The European Alliance
on Agricultural Knowledge for Development

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

The study benefited from the funding of the European Union

EIARD
European Initiative for Agricultural Research for Development

Impact assessment policies and practices of EIARD members

STUDY

agrinatura
The European Alliance on Agricultural Knowledge for Development
EEIG/GEIE

UNIVERSITY of GREENWICH | Natural Resources Institute

AGRINATURA-EEIG Secretariat
42 rue Scheffer
F-75116 PARIS
FRANCE
Fax: +33.1.53.70.21.56
secretariat@agrinatura-eeig.eu
http://www.agrinatura.eu/

Disclaimer

This report presents the views and judgement of the study team for '*Impact assessment policies and practices of EIARD members*' and does not necessarily represent those of EIARD. (European Initiative for Agricultural Research for Development).

Graphics:
Laura Bonaiuti

The study benefited from the funding of the European Union

The European Alliance on Agricultural Knowledge for Development

Impact assessment policies and practices of EIARD members

STUDY

by

Adrienne Martin and Valerie Nelson

January 2012

Table of content

Acknowledgements

We thank Dr Barbara Adolph and Dr Philippe Petithuguenin for their support given during this assignment and their useful comments on the draft. Thanks also to Dr Gilles Saint-Martin for his insightful suggestions and comments and to the EIARD country representatives for providing information which contributed to this study and for their comments on the presentation of preliminary findings of this study at their meeting in Montpellier in December 2011.

Executive summary

The overall aim of this study was to 'review and compare the policies and practices of different EIARD members in impact assessment to increase relevance, uptake and coordination of efforts by and for EIARD members, stakeholders and policy-makers'. The report discusses current methodological advances and debates in impact assessment and the current practices of EIARD members, using information provided by their national contact points, data base and literature searches and selected case studies, in order to develop practical recommendations for improvement and greater coordination and alignment.

Responses to an information request were received from 16 countries, plus the EC and relevant documents sent. However, there was limited information on impact assessment methodologies or the way in which findings had been used. Databases did not always provide links to relevant documents nor did 'evaluation' appear as a category in the thematic search options. A list of 224 projects was screened to select those which explicitly considered outcomes and impacts. 44 project or programmatic reviews were considered in detail, along with 6 synthetic reviews.

The terminology used in monitoring, evaluation and impact assessment is becoming increasingly complex as more importance is attached to these functions and their practice becomes specialised. There are different understandings of impact evaluation among research practitioners, evaluation specialists and also among EIARD members. According to the OECD-DAC definition, impact is defined as 'the positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended'. However, the more recent usages of 'impact evaluation' are specifically concerned with attribution of change to a programme or intervention, focusing on the question of what would have been the situation if the intervention had not been undertaken. This requires rigorous study designs in order to measure the net change in outcomes for particular groups of people. In contrast, there are more actor oriented and participatory approaches and techniques to map the logic of impact and then to assess it. The challenge is to select the best methods and combinations of methods for the purpose and resources available.

The EIARD strategy 2009-13 defines Agricultural Research for Development (ARD) as '*multi-dimensional in addressing the agricultural development challenges of developing and emerging economy countries*'. This definition is very broad and has implications for the types of material used as evidence for this study. For ARD it is important to assess the actual outcomes of the application of the research results, recognising that research makes its contribution within an 'innovation system' which requires partnerships with other types of organisations.

In 2000, the EIARD Task Force on Impact Assessment and Evaluation specified four main objectives of impact evaluations and identified elements of good practice which should be incorporated in their design. These elements helped to identify the criteria for analysing the evaluation and impact assessment reports from member countries. They included the location and context, scale and scope, funding and commissioning relationships, the purpose, objectives and types of impact assessment/evaluation; impact pathways; design, methods and tools; communication and dissemination.

The EIARD member countries have different relationships to evaluation and impact assessments, depending on whether the majority of their ARD funding is provided through direct support to programmes and projects or through multilateral organisations, such as the CGIAR. The majority of EIARD members combine the two strategies. Those providing unrestricted funding to the CGIAR as well as some of those giving restricted funds, rely solely on the impact assessment procedures within the CGIAR system, while others combine the system level reporting with their own evaluations. However, it was not always easy to establish from the reports analysed whether the evaluation was commissioned from sources independent of the funding body or research organisations.

The recent reform process within the CGIAR has established a new Consortium structure and proposed new arrangements for evaluation and impact assessment. It is recognised that a broader array of impact evaluation approaches is needed, with a focus on the contribution of research to poverty eradication, food security, gender equality and environmental sustainability. However, it will be a challenge for evaluators to balance these different requirements and to develop a suite of methods appropriate for a specific evaluation. Furthermore, the ideal is that evaluations should be coordinated and that all CGIAR funders would rely on a common results-based monitoring and evaluation framework.

The findings relating to evaluations commissioned by EIARD members, suggest that specification of the subject matter of the research and its contextualisation, the objectives and scope of the evaluation and discussion of other influencing factors, are well covered in current practice. However, other important recommendations are followed only in a minority of cases. These include, making explicit the model or concept of innovation, explaining the logic model underlying the programme or project and providing a statement or hypothesis of the impact that is expected. Only a few programme/multi-country evaluations specified impact pathways or results chains with clear linkages. It is clear that impact pathway analysis has not been a widely used tool.

There were few impact assessments involving measurement of actual changes and attribution. The majority were actually *outcome evaluations*, taking a broad 'plausible' linkages approach to examine the case for attributing change to the research intervention. They did not attempt to measure or attribute *impact*. Generally few of the studies appear to draw on recent developments in debates on rigorous impact assessment or alternatively, on participatory or narrative methods exploring a theory of change.

The small number of rigorous impact assessments analysed generally assessed impact at household level. The EIARD task force recommendation to explore the complex social, economic, political and institutional dimensions of impact requires complementary approaches and evaluation plans which combine different and complementary skill sets. Impact assessment in the sense of measuring attribution, utilising rigorous and statistically sophisticated methods is a specialised function and may be best contracted independently. The systematic application of qualitative methods also requires specialist expertise. This could be encouraged through more collective commitment (donors and national governments) to better coordination and joint funding of impact evaluations and for governments and agencies to reinforce efforts to generate exchange and apply knowledge from impact evaluations. Better collection and utilisation of monitoring data would also be helpful, since an understanding of the processes of project delivery, changing relationships and stakeholders' perspectives is important in interpreting results from impact studies.

Although not explicit in most documents, the recommendation that a plan for impact assessment and evaluation should be prepared before the project commences and be an integral part of project implementation does not appear to have been implemented. There is also scope for much wider inclusion of critical review and comment from different stakeholders, partners and beneficiaries.

The majority of the evaluations consulted made little reference to innovation as a multi stakeholder process. The recommendation of the EIARD task force to include a model or concept of innovation in evaluations, appears to have far to go and there remain important requirements for understanding and learning about the institutional context of agricultural research and development processes.

There was generally a lack of disaggregation of data in evaluations to indicate the impacts experienced by particular social groups. In particular there was limited identification of gender and poverty related impacts.

Considering the high level of direct support to the CGIAR, there was relatively little engagement in evaluation processes or utilisation of the CGIAR impact assessments. Only the EC appears to have examined the impact of CGIAR outputs. It is not clear whether other countries funding CGIAR directly had used these reports to inform their own decision making.

Few of the studies indicate in their methodology sections how they will seek to disseminate the findings to different audiences. The users of the findings of impact evaluations and the channels through which they will be reached are not well defined. Similarly, there was little information available on the ways in which impact assessments have been used internally and externally. Further understanding of how the findings and recommendations actually shaped policy and practice would require in-depth country case studies and face to face meetings. It was difficult to find policy briefs and summaries associated with any of the impact studies. With the exception of knowledge sharing, the 'process' uses of evaluations were not mentioned.

There are a number of recommendations which could enhance the efficiency and effectiveness of evaluation and impact assessment among EIARD members;

- Good evaluation and impact assessment begin with project design. It is important to develop impact oriented thinking, and to encourage the inclusion of evaluation plans and IA design in the project design and implementation plans. Focused baseline information collection can greatly enhance the capacity to assess outcomes and impacts.
- More emphasis on effective monitoring could encourage understanding of processes and achievements of ARD projects as they emerge. EIARD should engage with and support the current CGIAR efforts to develop more robust monitoring and evaluation systems in addition to the formal external ex post impact assessments conducted.
- There is a need to build understanding amongst those commissioning evaluations of ARD of the different kinds of evaluation and impact assessment and to guide choices in design and methods to be appropriate for specific objectives and circumstances.
- In commissioning evaluations, the expectations and type of evaluation required should be made clear. Terms of reference need to clearly specify the purpose of the evaluation and what is actually required. This is the basis for determining choice of methods.
- The planning of an evaluation should include a clear timeline, a step for critical review and comment from different stakeholders and a plan of action for communication of the evaluation findings.

- There is a need for development and agreement on procedures to encourage the sharing and dissemination of evaluation findings among EIARD members and their wider stakeholders. To help harmonize consistency and quality of reporting for ARD evaluations a best practice guide on quality standards specifically for ARD could be developed for EIARD members.
- Improvements to the ARD databases could increase the accessibility of evaluation and impact assessment reports. Evaluation reports and impact assessments should be categorised more clearly. Data bases could incorporate 'evaluation' as a search theme; include fields on dissemination of the evaluation findings; provide URLs to share reports on the website, and enhance the narrative descriptions of what has been done.
- Apart from improvements to existing data bases, EIARD members should explore their joint willingness to establish a web site or web page for open sharing of evaluation reports.
- Greater interest and commitment to develop joint studies should be encouraged to enhance methodological rigour and shared learning.
- There is a need to develop guidance for impact evaluation planning which helps in the selection of evaluation approaches appropriate for complex situations. The specific tools and techniques used should be consistent with the principles underpinning the evaluation and its objectives and tailored to facilitate exploration of the evaluation questions within the time and resources available.
- Multiple methods are preferable, exploring both the meaning and the measurement of project impacts. There is scope to innovate and support participatory, qualitative and mixed-methods, combining and sequencing different approaches and tools in evaluation.
- The development and use of flexible and non-linear programme theories of change should be incorporated as a standard tool within evaluation and specifically required in terms of reference. These take into consideration other actors and processes often neglected by logframes and linear impact pathways.
- The impact pathways should seek to disaggregate impacts for different stakeholder groups and in particular identify gender and poverty related impacts.
- Rigorous and quasi experimental approaches can be useful for assessing impact of specific sub-components of projects, particularly for technology components. They are less suitable for the complex, interactive, multi-stakeholder approaches of ARD.

1 Introduction

1.1 Context and objectives of the study

This study has been commissioned by the European Initiative for Agricultural Research for Development (EIARD)[1]. EIARD is being supported under the European Union funded Food Security Thematic Programme (FSTP) with the overall purpose of achieving 'coherent, coordinated, relevant and effective European policies for and investments in agricultural research for development that support the food security agenda'. Part of this support is for the production of three studies (of which this is one), and four policy briefs (one of which will be developed from this study).

The overall aim of this study is to '*review and compare the policies and practices of different EIARD members in impact assessment (IA) to increase relevance, uptake and coordination of efforts by and for EIARD members, stakeholders and policy-makers*'.

This relates closely to EIARD's goal which is to promote and implement coherent European policies at international, regional and sub-regional levels in order to increase the impact of agricultural research and development (ARD) on poverty reduction, food security and sustainable management of natural resources in developing countries.

According to the ToR, the study will:

- Contribute to ensuring that European ARD policies are based on knowledge of existing approaches and strategies, and related opportunities and challenges, making use of the latest available scientific and policy analysis;
- Analyse existing ARD policies, strategies, investments and programmes of EIARD members in relation to impact assessment and the extent to which they have factored in existing knowledge;
- Identify areas where increased coordination and harmonisation would be of benefit, and point out gaps, e.g. where policies are not based on latest evidence;
- Provide a basis for improvement and alignment of ARD policies, strategies and programmes so that they are able to more effectively respond to the challenges ahead;

A number of key study questions have been developed by the study team, based on the terms of reference;

1. What are the current paradigms and formal approaches to impact assessment among EIARD members? What are the commonalities and differences in approaches? How do these approaches compare with and draw on recent methodological developments in IA for ARD internationally? (e.g. in universities, 3ie, CGIAR etc.)

[1] Participation in EIARD is open to all European countries, members of the European Union or not, and to the European Commission. Active participants in 2011 were: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK and the EC (DG RTD, DG DEVCO). The study covers all the 27 Member States, plus the 2 other European countries and EC departments (30 entities).

2. What are the commonalities and differences in actual impact assessment methods for ARD (based on a selection of case studies) and in levels of investment? Are there differences between stated approaches and the actual methods of impact assessment?
3. How do the approaches and methods compare with the objectives of impact assessments and evaluations outlined in the EIARD task force paper "Impact Assessment and Evaluation in Agricultural Research for Development" (2001)? Do they explore complex social, economic, political and institutional dimensions and what is the relative emphasis on rigorous attribution compared to demonstrating broader 'plausible' links between research investments and developmental impacts?
4. Does impact assessment contribute to accountability (to EIARD member country governments and public; and to governments and civil society organisations in recipient countries)?
5. Does impact assessment contribute to learning at different levels about what works/doesn't work in ARD approaches?
6. Are lessons and evidence from impact assessments communicated to different audiences, particularly policy makers, and if so, how are they used to influence decision making and future investment in ARD?
7. What are the prospects for/barriers to greater coordination and harmonisation in impact assessment?

1.2 Scope and focus of the study

The focus of the study is on the policy and practical dimensions of impact assessment of agricultural research for development (ARD) among EIARD members. The report makes reference to the current methodological advances and debates in impact assessment, but the analysis is centred on the current practice of EIARD members, identified through information provided by the national contact points and selected case studies, in order to develop practical recommendations for improvement and greater coordination and alignment.

The main stakeholders are European public authorities, in particular the ARD programme managers, financial managers and policy makers; the agricultural research for development community in Europe and internationally, including southern ARD stakeholders and the regional ARD Fora.

1.3 Structure of the report

Following this introduction, section 2 sets out the methodology for this study. It begins by describing the process of information collection and sources utilised, then discusses definitions of evaluation and impact assessment and agricultural research for development. Section three analyses EIARD donors' main policies and approaches relating to impact assessment of ARD. Section three also includes discussion of allocation of resources to impact assessment. Section 4 presents the case study analysis of EIARD members' practice of impact assessment. Section 5 discusses the utilisation of the findings from impact assessments. The challenges and opportunities of impact assessment for ARD for EIARD members are covered in section 6, and conclusions and recommendations are presented in section 7.

2 Methodology of the study

2.1 Sources of information for review of policy documentation and ARD impact assessments.

Study questions 1 and 2 rely on developing a good understanding of the current approaches to IA among EIARD members through documentation, web sites etc. and a review of current literature on impact assessment of ARD and impact assessment methods more generally. This will indicate the extent to which EIARD members' utilise current advances in IA methods and approaches. Questions 3-6 concern the ways in which information and learning from impact assessment is used by member countries in communication and decision making. Question 7 examines the potential for greater harmonisation.

A list of relevant information needs was drawn up covering ARD policies and strategies, types of ARD funded and funding levels and commissioning practice; documentation on approaches to evaluation and impact assessment to understand the range of methods and approaches used, the different definitions used by EIARD members, levels of stakeholder involvement and the different purposes, uses and users of impact assessments.

EIARD members' national contact points were sent a letter via email requesting information on their approaches to understanding and measuring results from investments in ARD; their funding and commissioning of ARD impact assessments in the last 5 years, the main methodologies used, who conducted the impact assessments and also information on how the impact assessments had been used internally and externally (appendix 1). They were also invited to send examples of evaluation studies and impact assessments of ARD projects they have commissioned in the last 5 years which could be used as case studies and any other relevant documentation. In addition to the information and materials sent by EIARD members, the EIARD member profiles and policy documentation and information on funding of ARD on the EIARD and ERA-ARD[2] web sites were accessed.

Internet searches were made to identify further case study materials. Databases search were the European Information System on Agricultural Research for Development (InfoSys+)[3]; the DFID R4D database; FISA information system for Agriculture and Food Research. Search terms were: **evaluation, impact, assessing impact, impact assessment** used singly and in combinations **with ARD, agriculture, research, and development**. Other websites with relevant material were 3ie and Research into Use programmes, as well as the OECD DAC[4] evaluation resource centre and donor countries development assistance web sites. Literature searches were conducted to identify recent advances in approaches and methods for impact assessment in agricultural research, and in international development more generally.

[2]Agricultural Research for Development Dimension of the European Research Area (ERA-ARD).
[3] http://www.infosysplus.org/ (05-04-2011). The aim of InfoSys+ is to improve "access to European web resources in the areas of agriculture, environment, forestry, fisheries, socio-economics, rural-transformation and many others, devoted towards development." The database includes metadata on organisations, projects, funding opportunities, experts, news and events in ARD, which are categorized according to ARD-themes, activities, geo-focus and other attributes.
[4] http://www.oecd.org/pages/0,2966,en_35038640_35039563_1_1_1_1_1,00.html

The study team encountered some difficulties in accessing relevant information. Contact details were available for national contact points of 22 of the 30 potential EIARD members. Responses were received from 16 countries, plus the EC. Very few respondents were able to separate out funding for impact assessment from overall funding for ARD, or to comment on methodologies or the way in which impact assessments had been used. However, several sent potential case study documents. The extent to which the Infosys+ database was inclusive of all relevant projects was also uncertain and there are no direct links to relevant documents; hence not all documents of apparent interest could be accessed. 'Evaluation' did not appear as a category in the thematic search option. When 'evaluation' was used in a text search many non relevant items were generated.

The database and website searches plus the documents received resulted in a combined list of 224 projects. On examination, some of those turned out not to be impact assessments or evaluations which considered outcomes and impacts in terms of their contribution to development and these were excluded. The remaining studies were analysed using the conceptual framework (Table 2) and the results are discussed in section four.

2.2 Definitions used in the study

The two terms used in the terms of reference for this study; Impact Assessment (I.A.) and Agricultural Research for Development (ARD) are both open to varied interpretations.

2.2.1 Monitoring, Evaluation and Impact assessment

The terminology associated with the broad topic of monitoring, evaluation and impact assessment is becoming increasingly complex as more importance and higher expectations are attached to these functions and there is more refinement of the associated concepts and methodologies. Increasingly there is a gap between M&E and impact specialists on the one hand, and agricultural research and development practitioners on the other, in terms of their understanding of the concepts of M&E and impact assessment. There are also differences in terminologies used by different EIARD members as well as different understandings of the same terms.

European donor practice has generally tended to follow the definitions presented by OECD-Development Assistance Committee (OECD-DAC). These are used in the considerable body of information and guidelines developed by the European Commission to support programme and project evaluation[5]. A reasonable degree of consensus exists for definitions of monitoring and evaluation, but it is around impact assessment or impact evaluation that the greatest divergences appear. With respect to monitoring, the following definition is useful;

> **Monitoring** A *continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing (development) intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds*[6].

Monitoring involves regular collection of information relating to the planned activities and objectives over the duration of the intervention. It often focuses on a number of pre-defined

[5] http://ec.europa.eu/europeaid/how/evaluation/index_en.htm
[6] OECD-DAC, Glossary of Key Terms in Evaluation and Results, Paris, 2002

indicators. Monitoring information can allow '*for the correction of any deviation from the operational objectives, and thus improve the performance of the programme as well as facilitate subsequent evaluation*'[7]. It helps to answer the question 'is the project doing things right?' Some organisations refer to 'process monitoring' which looks at the developing relationships within the research and development process for the purposes of on-going learning, often using more qualitative approaches and eliciting feedback from participants. Other terms, 'impact monitoring' and 'participatory impact monitoring' indicate a focus on collecting information which will contribute to the understanding of changes, rather than implementation details. Monitoring is usually an internal project management function. Monitoring is distinct from evaluation, based on the timing, levels of analysis and relative specificity of the study[8]. Whereas monitoring is a continuous function, evaluation is an analysis and assessment of the project or programme's achievements at a specific point of time[9]. In the case of ex ante evaluation, anticipated changes are considered before the start, but evaluations are more usually at the mid or end point of a project. The OECD-DAC definition is widely used;

> **Evaluation** is '*The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or programme*'[10].

Evaluation is concerned with the extent to which the project's higher-level objectives have been met. It looks beyond pre-defined indicators to examine unintended as well as intended results and draws on wider data sources. It is often, though not exclusively, carried out by external evaluators. This type of evaluation is sometimes described as 'summative' as it summarizes achievements up to a particular time. It is distinguished from 'formative' evaluation which focuses on programme management, processes and activities while they are forming or happening, in order to shape or improve them.

The OECD-DAC evaluation criteria – relevance, efficiency, effectiveness, sustainability and impact[11] – represent different dimensions or perspectives for assessing the performance of a project.

- **Relevance** – the extent to which the objectives of the development intervention are consistent with beneficiaries' needs and problems, country needs, global priorities and partners' and donors' policies; whether the objectives continue to be relevant.
- **Effectiveness** – the extent to which the objectives of the development intervention were achieved, or are expected to be achieved, taking into account their relative importance.
- **Efficiency** – examines how resources – inputs, funds, expertise, time – have been converted to results and whether the results were achieved at a reasonable cost.
- **Sustainability** – the extent to which the benefits from a development intervention continue after major development assistance has been completed and the probability of continued long-term benefits.

[7] European Commission. 2004. Evaluating EU Activities – A Practical Guide for the Commission Services. Brussels: European Commission. OECD.
[8] Roche, C, 1999, Impact Assessment for Development Agencies: Learning to value change. Oxford Oxfam/Novib.
[9] Sida. 2004. Looking Back, Moving Forward. Sida Evaluation Manual. Stockholm: Swedish International Development Cooperation Agency (Sida). http://www.alnap.org/pool/files/evaluation_manual_sida.pdf (accessed
[10] OECD-DAC, Glossary of Key Terms in Evaluation and Results, Paris, 2002.
[11] To these, the EC add coherence/complementarity and community/value added.

- Impact - the positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended[6].

The OECD DAC criteria are utilized, with minor modifications, by many donors and development organizations, including EIARD members[12] Additional criteria used by some donors include *coherence* of the intervention vis a vis other development interventions and policies; *coverage*, which examines which groups are included in/excluded from a programme, and the differential impact on those included and excluded. Related concepts include equity (including gender equity and disability) and social exclusion, institutional development and partnerships; and *coordination*, which involves assessing harmonisation with other aid agencies and alignment with country priorities and systems.

'Impact' is distinguished by its long term rather than short or medium term nature. Impact is concerned with the 'big picture' changes in economic, environmental and social conditions that a project is working toward[13]. It also considers external influences and events[8]. Figure 1 shows the 'impact pathway', the relationship in simplified linear form, between inputs of resources which facilitate research activities, leading to the delivery of outputs, and the realisation of outcomes and impacts. Some sources refer to 'results' rather than 'outcomes', but the latter term more clearly differentiates outcomes from the outputs that produce them. It is widely recognised that the 'impact pathway' is often neither simple nor linear.

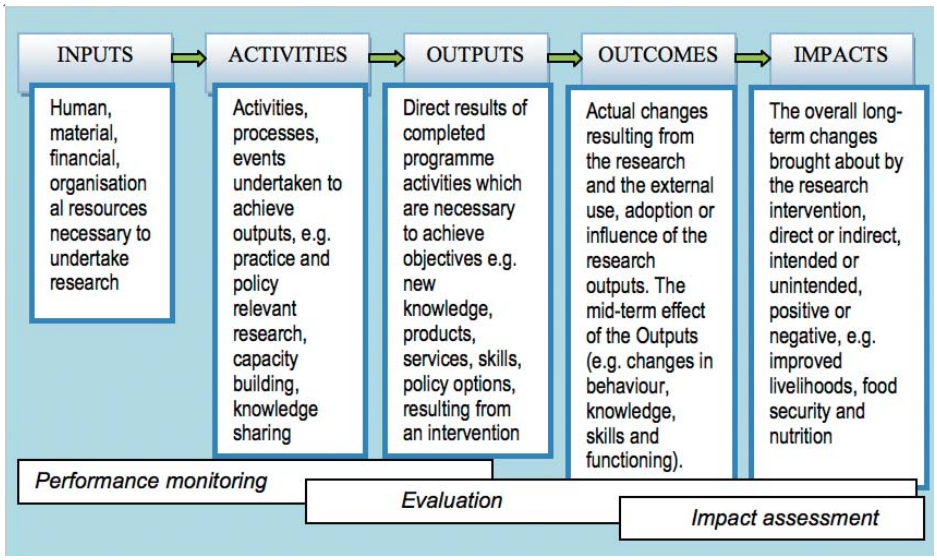

Figure 1 - Diagram of an impact pathway

[12] Examples are GTZ 2006. Working for sustainable results. Evaluation at GTZ. Eschborn: GTZ. http://www2.gtz.de/dokumente/bib/06-0796.pdf (accessed 3 January 2012).
GTZ 2007. Results Monitoring 2007: Evaluation Report on the Work of GTZ and Its Partners. 10th Cross-section Analysis 2006–2007. Eschborn: GTZ. http://www.gtz.de/en/dokumente/Evaluation-Report-0801.pdf (accessed 3 January2012). NORAD. 2006. Evaluation Policy 2006–2010. Part 1 Strategic priorities, Part 2 Evaluation Programme 2006–2008, Part 3 Guidelines for Evaluation of Norwegian Development Cooperation. Oslo: NORAD. http://www.norad.no/en/evaluation/handbook-and-reference-documents/_attachment/106229?=true&_ts=11eb62db87a (accessed 3 January2012). DANIDA 2006 Evaluation Guidelines. Copenhagen: DANIDA. http://www.netpublikationer.dk/um/7571/ (accessed 3 January 2012).
[13] T. Walker, M. Maredia, T. Kelley, R. La Rovere, D. Templeton, G. Thiele, and B. Douthwaite (2008) Strategic Guidance for Ex Post Impact Assessment of Agricultural Research. Report prepared for the Standing Panel on Impact Assessment, CGIAR Science Council. Science Council Secretariat: Rome, Italy ftp://ftp.fao.org/docrep/fao/011/i0276e/i0276e.pdf

The main source of confusion in interpretation of these terms relates to differences in what is actually meant by impact evaluation. The term 'impact assessment' is also used in different ways among different agencies.

Evaluation as described above assesses the effects or changes brought about by a particular programme intervention or policy. The dimension of impact is associated with understanding the longer term developmental effects of these interventions. Evaluations may involve wide stakeholder consultation and participatory evaluation by beneficiaries and other relevant actors to substantiate the nature and extent of the outcomes and impacts.

There is a very important distinction between this conception, and the more recent usage of **impact evaluation**[14] which is understood to have a more specific concern with attribution of change to a programme or intervention, focusing on the question of what would have been the situation *if the intervention had not been undertaken*[15]. Designs of such studies involve a 'counterfactual', based on a comparison between situations 'with' and 'without' the intervention in order to isolate the effect of the intervention and hence attribute impact. This type of rigorous impact evaluation is defined as '*analyses that measure the net change in outcomes for a particular group of people that can be attributed to a specific program using the best methodology available, feasible and appropriate to the evaluation question that is being investigated and to the specific context*'[16].

In this view, impact evaluation is considered as a specific approach within the larger toolkit of monitoring and evaluation (including broad programme evaluations, process evaluations, ex ante studies, etc.[17]

Impact assessment is sometimes used as an alternative term for impact evaluation, for example, by the Standing Panel on Impact Assessment of the CGIAR Independent Science and Partnership Council (ISPC). However they refer to it as ex-post impact assessment[18].

The term impact assessment is used in other contexts to refer to *ex ante* assessments carried out during project preparation. For example, the EC conducts impact assessments[19] of the potential economic, social and environmental consequences of new initiatives which provide evidence for decision makers on the advantages and disadvantages of possible policy options. Types of *ex ante* impact assessment include poverty impact assessment, social impact assessment and environmental impact assessment.

[14] White, H. 2009a. Some Reflections on Current Debates in Impact Evaluation. International Initiative for Impact Evaluation, Working paper 1. New Delhi: (3ie). http://www.3ieimpact.org/admin/pdfs_papers/11.pdf (accessed 6 June 2011).
[15] Impact Evaluation experience of the Independent Evaluation Group of the World Bank. http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/35BC420995BF58F8852571E00068C6BD/$file/impact_evaluation.pdf (accessed 18 September 2011). White, 2009a, Some Reflections on Current Debates in Impact Evaluation. International Initiative for Impact Evaluation, Working paper 1. New Delhi: International Initiative for Impact Evaluation (3ie). http://www.3ieimpact.org/admin/pdfs_papers/11.pdf (accessed 6 June 2011).
[16] 3ie Principles for Impact evaluation.
[17] http://www.3ieimpact.org/doc/principles%20for%20impact%20evaluation.pdf
[18] Leeuw and Vaessen, 2009. Impact Evaluations and Development. Nonie Guidance on Impact Evaluation. NONIE - The Network of Networks on Impact Evaluation.
T. Walker, M. Maredia, T. Kelley, R. La Rovere, D. Templeton, G. Thiele, and B. Douthwaite (2008) Strategic Guidance for Ex Post Impact Assessment of Agricultural Research. Report prepared for the Standing Panel on Impact Assessment, CGIAR Science Council. Science Council Secretariat: Rome, Italy ftp://ftp.fao.org/docrep/fao/011/i0276e/i0276e.pdf
[19] http://ec.europa.eu/governance/impact/index_en.htm

With reference to the definitions discussed above, this study is examining EIARD members' policy and practice in ex post impact evaluation and assessment of impact. It is not addressing ex ante evaluation or formative evaluation. The distinction has also been drawn in this study between evaluating outcomes and evaluating impacts: the former (evaluating outcomes) is drawn from monitoring data and systematic stakeholder feedback evidence and the latter (evaluating impacts) is based on impact evaluations and ex post impact assessments that incorporate a counterfactual and therefore allow specific comparisons to isolate the effect of the intervention and hence attribution of impact.

2.2.2 Agricultural Research for Development (ARD)

The EIARD strategy 2009-13 defines ARD as '*multi-dimensional in addressing the agricultural development challenges of developing and emerging economy countries (DEEC). The agricultural domain includes crop production and animal husbandry, agro-forestry, fisheries and aquaculture, food, agribusiness and related enterprises, as well as the sustainable management of the natural resources on which farming depends, the animal and human health related issues, and the socio-cultural and bio-diverse landscapes, food systems and ecologies in which it is embedded*'[20].

The EIARD strategy further notes that ARD is closely linked with other research sectors and themes, such as health, energy and environment, as well as social and institutional issues, such as gender and capability development. ARD provides technological, economic and institutional knowledge and innovations contributing to sustainable development. It encompasses public and private sector research, aiming at producing national and international public goods. Compared with the definition in the previous strategy[21] this formulation includes the concept of multi-dimensionality. ARD is now expected to broaden its agenda towards challenges of mutual interest to developing, emerging and industrialised countries. The dimensions and scope of ARD are defined as;

- ***fundamental and applied*** – dealing with upstream and problem-solving research;
- comprehensive – dealing potentially with research in any field and at any relevant scale, encompassing a wide range of scientific disciplines;
- ***multi-stakeholder*** –requiring iterative and inter-active loops of participatory diagnosis-to research-product processes that include all players and activities of the local innovation systems.
- ***international*** –carried out in and/or for developing and emerging economy countries, and in most cases with Southern ARD partners and International Agricultural Research Centres.
- ***global*** – as similar problems are widely shared among countries and as local interactions with world problems result from globalizations of all kinds;
- ***multiple policy*** purposed – because it contributes to various and different policies in different sectors.

The definition of ARD is very broad which has implications for the types of material used as evidence for this study, since impact assessments of any agricultural research relevant to DEECs could be considered. Nevertheless, this broad definition is within the context of the strategy

[20] EIARD Strategy 2009 – 2013. November 2008, p.3
[21] A strategy for the European Initiative for Agricultural Research for Development (EIARD 1) 2005-2010, summary, p1.

statements that ARD should '*provide the necessary understanding of rural development situations, of the drivers and impediments for sustainable rural development, and the required knowledge and innovations for the development of smallholders' livelihoods*'. Research into the broader and enabling context of rural development is needed in order to improve the impact of ARD on development.

European support to ARD seeks to promote partnerships, equity and balanced management responsibilities among partners in the South and in Europe and to increase national and continental capacities to plan and execute ARD activities. The strategy notes that national partners and research users must be involved in the formulation of research priorities, plans, implementation processes and *in assessment of the results and their usefulness*. It states that an Innovation System Approach will be systematically encouraged.

The EC Guidelines on ARD[22] produced in the same year as the strategy state that support to ARD '*should target those countries and regions where the incidence and depth of hunger demand urgent actions, which often correspond to remote, risk-prone and marginal areas. In terms of beneficiaries groups these should be smallholders farmers, and in particular women, who play an extremely important role in the agricultural systems of the poorest areas*'.

It also mentions a demand-driven and innovation system approach, the need for better linkages among researchers, farmers, agricultural advisory services and other stakeholders, including consumers, those in markets, agro-industry and academic institutions; to address all levels of the supply chain; to assist national agricultural research systems to implement changes in organisational culture, structure and systems; and to integrate principles of gender equality. This understanding of ARD has important implications for impact evaluation since it suggests the need to examine the contribution of research within the 'innovation system' (Figure 2[23]).

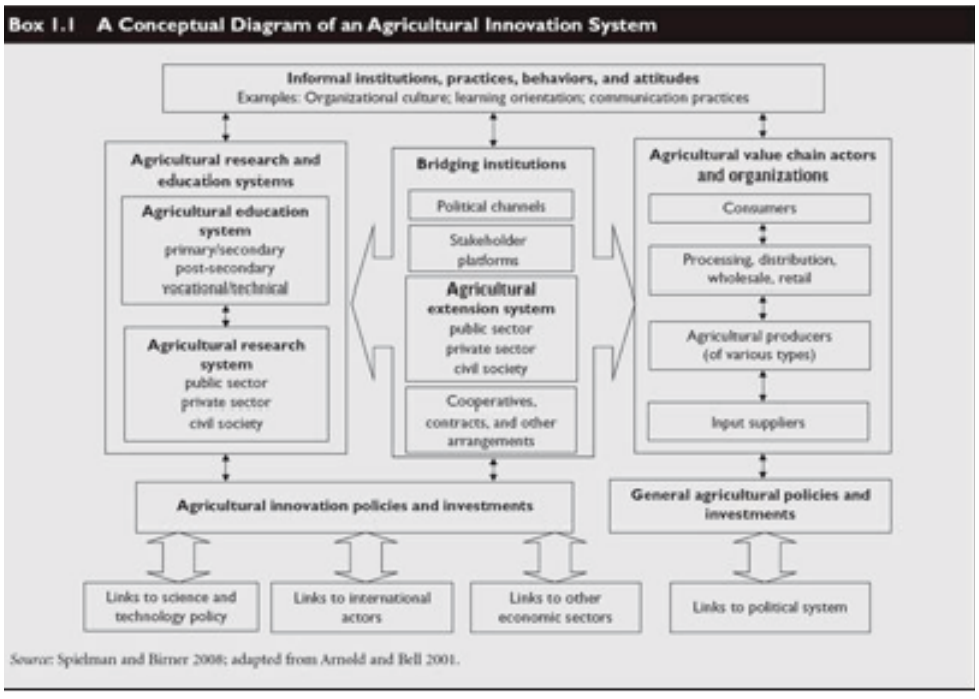

Figure 2 - A conceptual diagram of an Agricultural Innovation System

[22] European Commission Non paper Guidelines on Agricultural Research for Development, 2008.
[23] Spielman, D. and R. Birner. 2008. "How Innovative Is Your Agriculture? Using Innovation Indicators and Benchmarks to Strengthen National and Agricultural Innovation Systems." ARD Working Paper #41. Washington, DC: World Bank. http://knowledge.cta.int/en/Dossiers/Demanding-Innovation/Innovation-systems/Articles/Promoting-Agricultural-Innovation-Systems-Approach-The-Way-Forward

In the context of this broader conceptualisation of the nature and purpose of ARD, evaluations go beyond the assessment of technical research quality and technology performance; for ARD there is the additional requirement to assess its usefulness, in other words, the actual *outcomes* of the application of the research results. The complex linkages within the agricultural innovation system shown in figure 2 suggest the diversity of types of potential outcomes which should be explored in an evaluation; for example, on agricultural knowledge management, on institutions and policies, on public and private sector agriculture as well as specific outcomes for value chain actors, including farmers of different types and consumers. This perspective is well illustrated in EIARD thinking on impact assessment discussed in the next section.

2.3 EIARD and impact assessment

EIARD has from its early years emphasised the importance of impact assessment as a key factor in ARD. An EIARD Task Force on Impact Assessment and Evaluation was established in 2000 to develop a statement that could help to make impact assessment and evaluation studies more useful for investors in agricultural research and for the scientific community. Their paper, "Impact Assessment and Evaluation in Agricultural Research for Development" is used in this report as a basis for comparison with current EIARD members' policy and approaches to Impact Assessment[24].

The Task Force paper specified four main objectives of impact assessments and evaluations for EIARD:

a) To enhance the developmental impact of agricultural research investments for poor people;
b) To provide information on the returns to investments in agricultural research for development;
c) To derive strategic and programmatic lessons for future investments in agricultural research for development;
d) To provide information for use in public awareness work.

In relation to ***developmental impact***, the Task force considered that the primary motivation for undertaking impact assessments and evaluations should be to enhance the probability that investments in agricultural research will improve the livelihoods of poor people. They emphasise the importance of impact assessments and evaluations recognising and addressing the complex social, economic and political dimensions of pro-poor innovation. This argues against narrow interpretations of impact which focus on what is easily measurable. The paper suggests that evaluations of agricultural research for development should be designed and carried out within a holistic livelihoods framework, considering how research products and services have been used to address the range of people's needs and the developmental outcomes in terms of poverty, food security, natural resources use and conservation and sustainability.

The second objective of impact assessment meets the requirement of investors in research for information on the overall ***return on their investment***. The Task Force's paper recognised the difficulty and expense of attributing developmental impact to specific research investments and results. They argue that changes often result from dynamic, interactive, non-linear, and generally uncertain processes of innovation and lead to incremental changes over time and therefore it is

[24] Epilogue (2007) Impact Assessment and Evaluation in Agricultural Research for Development. Task Force on Impact Assessment and Evaluation, European Initiative for Agricultural Research for Development (EIARD). http://www.eiard.org/media/uploads/File/documents/impact_assessment_paper-1(1).pdf

seldom possible to identify clear cause-and-effect relationships between a given research project or its outputs and changes observed on the ground. It is often difficult to estimate a realistic counterfactual or scenario of what would have happened without the research. However, it is argued that it is possible to show plausible associations between research investments, outputs and development impacts, and to rule out rival hypotheses. This can be done by tracing out impact pathways (through assessing intermediate processes, products and outcomes that are requisites for impact) and establishing plausible links between the research investment and the observed development impact(s). Determining plausibility along impact pathways is more feasible than proving impact, and it provides more meaningful information for learning, programme steering, and accountability. Furthermore, it was suggested that evaluating the benefits and costs arising from investment in an agricultural research portfolio is more appropriate than single projects, as a few big successes will compensate for many failures. Evaluation of the whole portfolio would provide more useful information for investment decisions and policy making.

The third objective, to derive ***strategic and programmatic lessons*** to aid decision making and design of future investments in agricultural research for development, can be fostered through the open discussion of evaluation results and willingness to learn from failures. The Task Force point out that this implies a shift of focus from ex-post evaluations, where impact is often assessed long after the end of the project, towards more on-going, participatory assessments integrated into research programmes, which allow participants to learn and adjust the programme's direction while it is still active.

The fourth objective, also concerned with the use of information from impact assessments and evaluations, highlights their importance in providing information, for ***public awareness*** activities. These require convincing arguments on exactly how research has contributed to development goals. In using information from impact assessments for decision making, policy design and public awareness, it is important that information is tailored for different stakeholder interests and uses.

The paper concludes that scientific rigor is necessary, but it is not sufficient to meet the demands for developmental information and insight that diverse stakeholders expect from impact assessments and evaluations. It considers that searching for plausibility rather than proof of impact can help to produce useful information and insight at reasonable cost, while doing justice to the complexity of research-based innovation and encouraging well-grounded arguments and reasoned debate.

The Task Force paper identified elements of good practice in impact assessment and evaluation which should be incorporated in design. These are;

- Identification of the agricultural research investment, and a description of its context
- The model or concept of innovation
- The objectives, scope and limitations of the evaluation
- The logic model underlying the project or programme
- The statement and testing of an impact hypothesis
- A discussion of other factors that could have affected the observed changes
- A critical review and comment

These elements are discussed below (section 4) in the context of the criteria for reviewing the evidence base of recent evaluations and impact assessments from EIARD members.

2.4 Current initiatives and debates in impact evaluation

Concern with delivering impact is at the forefront of discussions on agricultural research for development (ARD). The report by the Global authors' team for the Global Conference on Agricultural Research for Development (GCARD) 2010, emphasised the importance of translating agricultural research into wide scale, sustainable development impacts on food security, poverty and the environment[25]. Yet the practice of evaluation and impact assessment has not always provided the required evidence, nor influenced policy making and professional practice. More broadly in development circles, there is a resurgence of interest in evaluation and impact assessment, linked to a renewed emphasis on value for money. It is consistent with the emphasis on measuring progress and development effectiveness articulated in the Paris declaration (OECD/DAC Working Party on Aid Effectiveness 2008).

Since the EIARD Task Force paper discussed above, there have been a number of initiatives focusing on impact assessment and evaluation of agricultural research for development as well as initiatives addressing issues of impact assessment across sectors. There have been considerable advances in the theory and practice of impact evaluation. This section reviews some of the current approaches to impact assessment internationally and the associated debates.

2.4.1 Impact evaluation and the CGIAR

The Consultative Group on International Agricultural Research's (CGIAR's) Standing Panel on Impact Assessment (SPIA) has been working since 1995 to enhance the quality and methodological rigour of ex post impact assessments and to encourage feedback into research planning. The CGIAR impact web site (http://impact.cgiar.org) has many examples of impact assessments and papers on methodology and guidelines[26] for conducting impact assessments. Their focus is on developing rigorous methods for assessing the impacts of agricultural research and technology development on poverty, hunger and food security, both directly and indirectly. Two broad types of ex post impact assessment are outlined; i) economic rate of return assessments, either aggregate, or on specific technologies or policies generated by investment in agricultural research (these form the majority of studies included in recent reviews of the impact of CGIAR-related agricultural research[27]); and ii) multi-dimensional impact assessments which variously explore the effects of technological change on growth, labour markets and migration, producer welfare and poverty. The latter focus mainly on larger or more visible technologies and make use of mixed methods with a view to promoting learning related to technology transfer, policy and economic development. Such assessments are more expensive than those focusing on the economic rate of return. They are conducted when longer term impacts of technology adoption are likely to have become evident, in contrast to the technology-focused economic assessment of costs and benefits at the adopter level.

A review of evaluations of CGIAR projects in 2003 noted that benefits compared to investment show high ratios in three areas, genetic improvement of rice and wheat and cassava mealybug

[25] (http://www.fao.org/docs/eims/upload//282426/GAT_Report_GCARD_2010_complete.pdf)
[26] Walker, T., Maredia, M., Kelley, T., La Rovere,R., Templeton, D., Thiele, G., and Douthwaite,B. 2008. Strategic Guidance for Ex Post Impact Assessment of Agricultural Research. Report prepared for the Standing Panel on Impact Assessment, CGIAR Science Council. Science, Council Secretariat: Rome, Italy.
[27] Maredia, M. and D. Raitzer 2006. CGIAR and NARS partner research in sub-Saharan Africa: evidence of impact to date, Science Council Secretariat, Rome.

biocontrol[28]. An update of this study in 2010 confirmed the strongly positive impacts from CGIAR research relative to investment. There have also been beneficial impacts from natural resources management and policy research, although these have tended to be locally and nationally rather than internationally. Crop genetic improvement research has had the most profound documented positive impacts[29]. The study recommended that the CGIAR prioritize impact assessment of resource management and policy research to deepen its understanding of the social and environmental impacts of its work.

The most recent review of ex post impact assessments within the CGIAR (Janvry *et al*, 2011[30]) argues for enhancing rigor in these assessments by making greater use of recently developed research designs and analytical tools. It identifies two main weaknesses in current approaches to microeconomic impact analysis - the problematic formation of the counterfactual non-adopting group against which to compare adopters, and failure to account for spillovers between adopters and non-adopters. The paper is critical of some of the methods currently used for assessing impacts and argues for increased use of experimental designs, including randomised controlled trials (RCTs)[31]. These are seen as important in eliminating selection bias between treatment and control groups. The paper acknowledges critiques of such approaches within the evaluation field, on grounds of cost, ethical dimensions, and their limited scale, however, considers that they are valuable in providing more rigorous assessment of research. It is interesting that the discussion focuses on the micro economic analysis of the impacts of adoption of new varieties, rather than more complex natural resource management technologies or policy research.

2.4.2 CGIAR reform and impact evaluation

The recent reform process within the CGIAR has constituted a new Consortium structure of CGIAR-sponsored Centres, and a CGIAR Fund managed by donors and partners. The new CGIAR Strategy and Results Framework (SRF) which was approved by the Fund council in April 2011, is aligned with three overarching themes: Food for People, Environment for People and Policies for People. It defines four system level outcomes[32] which the CGIAR research programmes will deliver.

- *Reduced rural poverty*. Improved productivity and better developed markets will deliver agricultural growth in which the rural poor participate;
- *Improved food security*. increasing supplies of key staples will buffer communities against price rises and volatility making food affordable for millions of poor people;
- *Improved nutrition and health.* Improved crop varieties and diversified production systems will provide the nutrients often lacking from the diets of poor people, particularly women and children;
- *Sustainably managed natural resources*. Only through this outcome can poor farmers benefit

[28] Raitzer DA. 2003. Benefit-Cost Meta-Analysis of Investment in the International Agricultural Research Centres of the CGIAR. CGIAR Science Council Secretariat, Washington, and FAO, Rome
[29] Renkow, M and D. Byerlee, 2010. The impacts of CGIAR research: A review of recent evidence. Food Policy Volume 35, Issue 5, October 2010, Pages 391-402
[30]de Janvry, Alain, Andrew Dustan, and Elisabeth Sadoulet. 2011. 'Recent Advances in Impact Analysis Methods for Ex-post Impact Assessments of Agricultural Technology: Options for the CGIAR' Report prepared for the CGIAR. April 2011 http://areweb.berkeley.edu/~sadoulet/papers/deJanvryetal2011.pdf
[31] An example of the use of randomised controls for impact assessment is the Sub Saharan Africa Challenge Programme - see page 49.
[32] CGIAR, Changing Agricultural Research in a Changing World.
http://www.cgiarfund.org/cgiarfund/sites/cgiarfund.org/files/Documents/PDF/CGIAR-SRF-%20Brief-March%202011_2_PAGER.pdf

from healthy ecosystems and sustain high-level agricultural productivity, particularly in the light of climate change.

The development of new Consortium programmes (CRPs, previously called Mega Programmes, MPs) has been accompanied by efforts to align monitoring, evaluation and impact assessment. The CRPs are multi-centre initiatives, globally or regionally based, requiring a strategic approach to achieve impact in one or more of the above system outcomes. Among other principles, they should integrate research across CGIAR core competencies and centres and engage with stakeholders and develop effective partnerships. Their research is conducted through an agricultural research for development (AR4D) approach. The principle is one of shared responsibility[33]. Programmes are seen as an instrument for greater alignment of research outputs with system level outcomes.

The CRPs include some with a crop or commodity focus; some with a systems base and some with a problem or policy focus. Fifteen programme areas have been identified. Three of these programmes (Forests Trees and Agroforestry - Livelihoods, Landscapes and Governance; Climate Change, Agriculture and Food Security; and the Global Rice Science Partnership) were approved in 2010. By November 2011 an additional six programmes[34] had been approved (or approved with light adjustments) by the Fund Council. A further five programmes[35] were approved with some conditions and one programme is under revision for resubmission (Integrated Systems for the Humid Tropics).

CRP6, Forests, Trees and Agroforestry: Livelihoods, Landscapes and Governance[36] is unusual in that its core research activities are embedded in specific impact pathways for each component designed to deliver distinct but interlinked outcomes which together will generate a common set of impacts. It is stated that the research will result in increased awareness and understanding among key stakeholders, practitioners and policymakers of the problems and opportunities for improving technical practices and developing *more appropriate and effective policies and governance mechanisms that deliver real-world impacts*. It is not yet clear the extent to which other CRPs will follow this example.

There has been considerable emphasis placed on the importance of delivery of impact from the new programmes, particularly by the Independent Science and Partnership Council (ISPC) of the CGIAR. Management for results is one of the four key principles of the new CGIAR. The SRF is intended to set '*common goals (in terms of development impacts) strategic objectives and results (in terms of outputs and outcomes) [that are] to be jointly achieved by the Fund, the Consortium and the bilateral funders to the Centres within a certain time frame*'[37] . Because the programmes are new, the

[33] CGIAR Strategic results framework. A strategy and results framework for the reformed CGIAR, July 2011 http://www.cgiarfund.org/cgiarfund/sites/cgiarfund.org/files/Documents/PDF/CGIAR-SRF-March%202011_BROCHURE.pdf
[34] Aquatic Agricultural Systems for the Poor and Vulnerable; Wheat-Global Alliance for Improving Food Security and the Livelihoods of the Resource poor in the Developing World; Maize- Global Alliance; Roots, Tubers and Bananas; More Meat, Milk and Fish by and for the Poor; Water, Land and Ecosystems.
[35] Policies, Institutions, and Markets to Strengthen Food Security and Incomes for the Rural Poor; Grain Legumes; Dryland cereals; Agriculture for Improved Nutrition and Health; Integrated Agricultural Production Systems for the Poor and Vulnerable in Dry Areas
[36] CGIAR Research Program No. 6: Forests, Trees and Agroforestry: Livelihoods, Landscapes and Governance. Executive summary. http://www.cifor.org/fileadmin/fileupload/crp6/CRP-ExecutiveSummary_en_web.pdf
[37] ISPC 2010, Note on the CGIAR Strategic Results Framework. http://www.sciencecouncil.cgiar.org/fileadmin/templates/ispc/documents/About_Us/Contributions_to_the_CGIAR_change/SC_ISPC_Note_on_SRF_for_the_e-conference.pdf

discussion to date has focused on ex ante assessment and the need for clear articulation of the pathways to delivery of outputs, outcomes and impacts. However, the unpredictable nature of research and the time lags before impact is manifest are also recognised. In discussing the need for independent evaluation to complement internally organised monitoring and evaluation, the ISPC states that 'the basic challenge of evaluation governance design consists in sustaining full independence *without incurring isolation*'[38]. This acknowledges the need for evaluators to have sufficient knowledge and understanding of the content and context of the research, whilst being fully independent of its implementation and having no vested interest in its outcomes. The impact assessment needs of the new system are evolving, and include the following, as defined by the CGIAR Science Council[39];

- Clarifying objectives of and priorities for ex post IA – accountability and learning. A broader array of evaluation approaches is needed, balancing the accountability function, with feedback and operational learning to improve the design (through ex ante IA) and conduct of future research efforts.
- Developing and applying new IA methods to document adoption, influence and impact in and address issues such as attribution and counterfactuals, especially in the more challenging areas of IA as NRM, policy and biodiversity research and training/capacity building, but also in the traditional areas for IA.
- Methodological advances in IA are needed, including testing and validating qualitative methods derived from the sociological and anthropological traditions in development studies, as complements to economics-based IA approaches that could form part of a multi-dimensional comprehensive approach to the study of impact.
- Broadening the scope of IA beyond partial (economic) assessments, advancing further down the impact pathway toward indicators that reflect more closely CGIAR goals. More comprehensive assessments of other dimensions of impact, relating to poverty, food and nutrition, security effects and environmental benefits/losses.
- Making ex post IA more utilisation focused, improving use for strategic feedback on research impacts; how far results of ex post IAs reflect the earlier projections from ex ante IAs; how spillovers can be maximised so that the returns to R & D investments are enhanced. Enhancing the coverage and rigour (credibility) of IA efforts to improve the consistency of IA coverage across the system.

On the broadening of the approach, the '*logic of CGIAR's monitoring of impacts on rural poverty shifts from a focus on understanding the impact of a particular technology on the incomes of the rural poor, to understanding the complex of factors required to significantly reduce poverty rates on some level*'[40]

The SPIA Strategy and Operational Plan 2011-13[41] acknowledges that this shift in focus responds to donors concerns that research investment should contribute to global development goals of

[28]Interim Independent Science and Partnership Council, Contribution to the discussion on Mega Program M&E and the independent evaluation arrangement in the new CGIAR
http://www.sciencecouncil.cgiar.org/fileadmin/templates/ispc/documents/About_Us/Contributions_to_the_CGIAR_change/SC_ME_of_MPs_in_the_new_CGIAR_Jan_2010.pdf
[39] CGIAR Science Council. 2009, Defining the role of an Independent Science and Partnership Council.
(A Discussion Document contributing to the CGIAR Transition)
http://www.sciencecouncil.cgiar.org/fileadmin/templates/ispc/documents/About_Us/Contributions_to_the_CGIAR_change/SC_Defining_the_role_of_an_ISPC_April_2009.pdf
[40]A strategy and results framework for the CGIAR, February 20, 2011.
http://www.cgiarfund.org/cgiarfund/sites/cgiarfund.org/files/Documents/PDF/srf_feb20_2011.pdf
[41] CGIAR Standing Panel on Impact Assessment: Strategy and Operational Plan 2011-13
http://impact.cgiar.org/sites/default/files/images/SPIAstrategy2011-13.pdf

poverty eradication, food security, gender equality and environmental sustainability. SPIA is seeking to address the methodological challenges of assessing these dimensions of impact as well as conducting evaluations on research areas less investigated in the past such as legume crop improvement and management research and germplasm collection, conservation, characterisation and evaluation. It will also put more emphasis on assessing impacts of research in policy, natural resources management and genetic resources. The overall impression from the strategy documents is that different approaches and emphases in impact assessment are all represented. It will be a challenge for evaluators to sort out the balance of these different requirements and to develop a suite of methods appropriate for a specific evaluation.

In addition to external evaluations of each centre (every 5 years), the Consortium board will commission external evaluations of the new programme components (every 4 years) and cross-cutting issues. These evaluations will contribute to the independent evaluations carried out every four years, commissioned by the Fund council with support from the Independent Evaluation Arrangement of the CGIAR (IEA)[42]. It is intended that the approach to evaluation should be coordinated; the joint declaration states that '*All CGIAR funders will rely on a common results-based monitoring and evaluation framework as part of the common operational framework*'.[43]

The thinking on impact assessment is still evolving. According to earlier CGIAR documents, the Independent Science and Partnership Council (ISPC) were to be responsible for ex ante evaluation of the programmes, foresight studies and ex post impact assessment through the Standing Panel on Impact Assessment (SPIA). The CGIAR Research Programmes (CRPs) and Centres will collaborate in impact assessments across the System Level Outcomes of the CGIAR Strategy and Results Framework. However, the policy for IEA states that ex post impact assessment is the responsibility of the CRPs, supported methodologically and for particular studies by the SPIA. According to SPIA's operational plan it will continue to encourage Centres to undertake quality impact assessments and aggregate these into estimates of system wide impacts. It will also seek to improve communications with stakeholders through the CGIAR impact website and disseminating of outputs in various forms, holding conferences, interacting with donors and networking with Centres' IA focal points.

There is potential for overlap or duplication in the scope of these different evaluation functions[44], conducted at different levels (projects, programme components, programmes) and commissioned by different bodies (Table 1). However, SPIA states that it will ensure that impact assessment work is complementary to and integrated with the work programme of the Independent Evaluation Arrangement (IEA)[45] .

2.4.3 The search for rigour in impact evaluation

There are some important cross-sectoral initiatives in impact assessment which have relevance for the discussion of impact assessment in ARD. Several of these are specifically concerned with developing rigorous impact evaluation methods, by which they mean methods which include analysis of attribution, establishing as far as possible a causal link between the intervention and specific impacts. Such approaches require the construction of a rigorous counterfactual (often

[42] CGIAR policy for Independent External Evaluation. CGIAR Fund, January 2012.
[43] CGIAR Joint Declaration, 2009, Annex Conduct, p8.
[44] Rothschild and Tollens, 2011 New Terms of Reference for Joint Review Missions applicable in 2011 and 2012 in the EIARD Framework for the Monitoring & Evaluation of the CGIAR.
[45] CGIAR Standing Panel on Impact Assessment: Strategy and Operational Plan 2011-13

Table 1 - 'Research to impact' pathway and focus of evaluation

Research Outputs	Research Outcomes	System Level Outcomes	Development impact
Change in knowledge Change in capacity Change in technology Change in materials Change in policy options Change in awareness/ understanding	Recognition/appreciation of research knowledge Use of knowledge by partners Mobilization of new capacity and resources Extension of technology/ Materials Change in policy environment	Change in action/behaviour of potential beneficiaries Change in productivity Change in equity/ Empowerment Change in market condition Change in investments Change in security of assets/ habitats	Changes in well-being, improved livelihoods, availability of ecosystem services, opportunities, choices, reduced risks
Monitoring by Consortium program management	Independent External evaluations (Consortium programme management)	Independent evaluations (Fund council/ IEA)	Ex post impact assessment (IEA/ SPIA?)

Source: Adapted from CGIAR, 2009. Voices for Change: The New CGIAR.

through the use of comparison groups) and where possible, the use of quantitative measures of impact to exclude contextual influences and establish causality of programme effects. As indicated in the context of the CGIAR, these involve the use of *experimental* or *quasi-experimental* designs. An experimental design involves the random assignment of individuals or households either as beneficiaries, or as a control group which does not receive the service or good being provided by the project. This is to ensure that the two groups are as similar as possible apart from the intervention under study. The approach is designed to show causal relationships between certain outcomes and the "treatments" or services aimed at producing these outcomes. It therefore requires control over who does or does not get the intervention and prevention of 'spillover' between the intervention and non intervention groups.

A quasi-experimental approach can be used where the intervention cannot be randomly assigned. A non intervention control group is constructed which matches as closely as possible the characteristics of the 'intervention' group. This is necessary to avoid *selection bias*— when those in the treatment group are different in some way from those in the control group. There are a number of techniques which help address this[46]. These methods, particularly the experimental methods, often involve the use of large scale sample surveys, and sophisticated statistical analysis. Examples of three initiatives in rigorous impact evaluation are discussed here.

The first is the International Initiative for Impact Evaluations, 3ie (http://www.3ieimpact.org/). Among its members are some EIARD members - Danish International Development Agency (DANIDA), Department for International Development (DFID), Irish Aid, the Netherlands

[46] Leeuw and Vaessen, 2009. NONIE Guidance on Impact Evaluation .The Network of Networks on Impact Evaluation.

Ministry of Foreign Affairs and the Swedish International Development Cooperation Agency (SIDA). 3ie is making an important contribution to methodological innovation in impact evaluation. It funds studies that are built around a credible counterfactual with an evaluation design based on the underlying programme theory to learn what works and why and at what cost. It emphasises the synthesis and dissemination of this evidence to help build a culture of evidence-based policy making (Box 1).

Box 1: Examples of 3ie studies and methods

Access, Adoption, and Diffusion: Understanding the Long-term Impacts of Improved Vegetable and Fish Technologies in Bangladesh
- *Difference-in Difference (DID), Propensity Score Matching (PSM*
Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya - Randomised Control Trials (RCT
The Impact of Improved Maize Varieties on Poverty in Mexico: *A Propensity Score-matching Approach*
http://www.3ieimpact.org/database_of_impact_evaluations.html

The Nonie Network of Networks on Impact Evaluation http://www.worldbank.org/ieg/nonie/ is concerned with promoting more and better impact evaluations among its members through sharing of methods and learning-by-doing. It has produced a useful guidance document[47] covering nine key issues in impact evaluation, including methodological and conceptual issues, including the scope and focus of the evaluation, the theory of impact, the attribution problem and mixed methods approaches. It also covers the management of impact evaluation addressing feasibility, benefits and costs. While the network is concerned with advancing rigorous impact evaluation, it recognises that this is more complex than advocating a single method or particular design. Choice of method depends on the questions and objectives of a particular impact evaluation and a mixed methods approach is advised.

Thirdly, the Abdul Lateef Jameel Poverty action lab (J-PAL) http://www.povertyactionlab.org is a network of users of Randomized Evaluations seeking to answer questions critical to poverty alleviation. J-PAL's mission is to reduce poverty by ensuring that policy is based on scientific evidence. The concerns are similar to those raised by Janvry et al, 2011 including discussion on the methods of randomisation, the level at which it should take place, what sample size should be chosen[48] and possible sources of bias such as spillover and cross over effects. These occur when the comparison group is influenced by the intervention, thus invalidating the comparison with the intervention group or when individuals in the control group find their way into the intervention group. Other considerations are whether the evaluation design is fair and ethical, politically acceptable and logistically feasible.

2.4.4 Addressing complexity in evaluations

Amongst those concerned with the developmental impacts of ARD, there is a growing perspective that multiple and complementary methods are needed to assess complex multiple processes of change (for example, as indicated in the EIARD task force paper summarised above).

[47] Leeuw and Vaessen, 2009. NONIE Guidance on Impact Evaluation .The Network of Networks on Impact Evaluation
[48] Duflo, Esther, Glennerster, Rachel, and Kremer, Michael, "Using Randomization in Development Economics Research: A Toolkit" (2006). MIT Department of Economics Working Paper No. 06-36.

Some argue that in ARD and processes of rural and agricultural change, there are far too many variables and possibilities emerging and interacting dynamically to rely on simple counterfactuals[49].

In contrast to the initiatives associating rigour in impact evaluation with experimental or quasi experimental approaches, there are others which utilise more actor oriented approaches. These include different techniques to construct the logic of impact and then to assess it – through for example, participatory development and assessment of indicators and evaluation criteria, use appreciative enquiry methods[50] and quantification of qualitative perceptions of change, Outcome Mapping[51] and the Most Significant Change (MSC) method[52]. These derive evidence of change and impact through processes of social verification and validation across different stakeholder groups. They have their own standards of rigour. Such methods are seen as particularly important where the desired changes take place within a complex of multi stakeholder relationships and responsibilities, and where changes in relationships, policies, attitudes, capacities and quality of communication are important objectives.

The Agriculture Learning and Impacts Network (ALINe) http://www.aline.org.uk is concerned with improving monitoring and evaluation in agricultural development projects. It promotes people-centred performance measurement that provides accountability to both beneficiaries and donors, arguing that participation can improve the relevance, efficiency and effectiveness of initiatives. The methodological innovations emphasised here are those which bring in farmers voices alongside those of other stakeholders, e.g. development of theories of change, participatory impact pathways and outcome mapping.

ILAC, the Institutional Learning And Change Initiative (an inter centre initiative of the CGIAR) has been working to improve the evaluation of collaborative agricultural research for development programmes so as to enhance the impact of research on poverty and other development goals. It deals with the difficult issues of the impact of ARD on poverty and the design of innovative and novel approaches to evaluation. It provides support to decision makers on the choice of the most appropriate combinations of evaluation approaches and methods[53] and their implementation. It is particularly concerned with exploring a wider range of positive and negative impacts connected with multi stakeholder partnerships in research; with building evaluation capacity and advocating for better evaluation policies and practices in ARD organizations. ILAC provides information on methodologies and techniques for different kinds of evaluations.

The Innovation Systems Framework also offers a useful framework for analysis, monitoring and evaluation of networks of organizations in research and innovation[54] and links localized networks to a broader enabling national and international policy and trade environment. Network

[49] Patton, M.Q. 2010. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. New York, NY: Guilford Press.
[50] http://appreciativeinquiry.case.edu/intro/definition.cfm. Preskill and Coghlan 2003.
[51] Earl, S., F. Carden, and T. Smutylo. 2001. Outcome Mapping: Building Learning and Reflection into Development Programs. Ottawa: IDRC. http://www.idrc.ca/en/ev-9330-201-1-DO_TOPIC.html (accessed 3 Jan 2012).
[52] Davies and Dart 2005. The most significant change (MSC) technique: A guide to its use. http://www.mande.co.uk/docs/MSCGuide.pdf (accessed 3 January 2012). MSC is particularly useful where criteria for evaluation have not been specified as part of a project design and M&E framework.
[53] http://www.cgiar-ilac.org/content/tools-and-methods-me
http://www.cgiar-ilac.org/content/evaluation-studies-and-reports
[54] Hall, A., V. Rasheed Sulaiman, N. Clark, and B. Yoganand. 2003. From measuring impact to learning institutional lessons: An innovation systems perspective on improving the management of international agricultural research. Agricultural Systems 78(2):213–241.
World Bank 2006 Enhancing Agricultural Innovation: How to Go Beyond the Strengthening of Research Systems. Washington: The International Bank for Reconstruction and Development/The World Bank. http://siteresources.worldbank.org/INTARD/Resources/Enhancing_Ag_Innovation.pdf (accessed 6 June 2011).

models (http://www.mande.co.uk/networkmodels.htm) are useful for the analysis of such situations with multiple actors (people and /or organisations).

Although theory based evaluation has been around for some time[55] there have been more recent discussion on developing 'theory of change' models to guide evaluations[56]. These can be conceptualised as a series of cause and effect relationships linking inputs and activities to expected outcomes and impacts and to overall goals. They address processes in more detail, generally involving an examination of the context, assumptions, and preconditions and presenting them in a more interlinked visual format.

The choice of methods and approaches depend on the context and purpose of an evaluation, what kinds of questions are being asked and what type of analysis is required. It is important therefore, to define the expectations of impact evaluation studies before engaging in their implementation. The challenge is to select the best methods and combinations of methods for the purpose and resources available. Some types of evaluation are more appropriate for particular kinds of intervention[57] . Among the considerations are;

- The acceptability of approaches to different client groups – donors and country partners; what is seen as constituting 'evidence' and the level of rigour required
- The ethical implications of 'excluding' potential beneficiaries from a development programme in order to construct a 'control' group.
- The main function and purpose of the evaluation – including the relative weight given to accountability or 'proving' impact, or to learning lessons for the future and 'improving' performance.
- The type of programme being evaluated and the degree of complexity (whether the programme has a clear objective with outcomes clearly linked to the intervention, and can be isolated, manipulated and measured, or whether it is a more complex initiatives such as holistic livelihoods approaches, research conducted with an innovation systems perspective or policy research.
- The scale of the evaluation, the feasibility and the resources and time required for different approaches.

2.5 Conceptual framework for analysing IA in ARD

The EIARD Task force paper outlined seven elements of good practice in impact assessment and evaluation. These form a major part of the criteria for analysing the evaluation and impact assessment reports from member countries. The criteria (summarised) are;

1. ***Identification of the object of the evaluation and its context***. A clear description of the research and the intended impact including the different actors in the innovation process.

[55] Weiss, C. H. 1995. Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families", in J. Connell, A. Kubisch, L. B. Schorr, and C. H. Weiss (eds) New Approaches to Evaluating Community Initiatives, New York: Aspen Institute.
[56] White, H. 2009a. Some Reflections on Current Debates in Impact Evaluation. International Initiative for Impact Evaluation, Working paper 1. New Delhi: International Initiative for Impact Evaluation (3ie). http://www.3ieimpact.org/admin/pdfs_papers/11.pdf.
[57] Duignan, 2009 Selecting impact/outcome evaluation designs: a decision-making table and checklist approach. Outcomes Theory Knowledge Base Article No. 256. http://knol.google.com/k/paul-duignan-phd/selecting-impactoutcome-evaluation/2m7zd68aaz774/115 (accessed 3 January 2012).

Listing and describing the activities and processes through which research produced its outputs and early outcomes to construct a chain of plausible impact.

2. ***Model or concept of innovation***. Attribution is difficult given the complex ways in which research impacts on the livelihoods of poor people. Making the model or concept of innovation explicit and superimposing the research strategy onto it, helps to understand the internal logic of the impact assessment and to check the completeness of the inquiry.
3. ***Objectives, scope and limitations of the evaluation to be clarified***, e.g. to enhance the developmental impact of ARD investments; to generate lessons from experience; providing information for project management, priority setting and planning; and justifying research investments. Indicate the limitations of the evaluation, in attempting to bridge the "attribution gap" between documented results and plausible impacts further down the impact pathway.
4. ***Logic model underlying the project or programme*** – what was hoped to achieve with the investment, how the research activities were meant to contribute to development objectives. A plan for impact assessment and evaluation should be prepared before the project commences and it should be an integral part of project implementation.
5. ***Statement and testing of an impact hypothesis***. The plausibility of IAE results can be enhanced if they have been explicitly developed in relation to an impact hypothesis - a statement about the impact that is expected to be found. This can often be derived from a logframe.
6. ***Discussion of other factors that could have affected the observed changes*** and their potential effects weighed.
7. ***Critical review and comment***. Plausibility and credibility are strengthened when dissenting points of view of different stakeholders, partners and beneficiaries are presented and discussed.

These elements, in combination with the Nonie evaluation guidelines, were considered in compiling the framework below (Table 2) for analysing the impact evaluations commissioned by EIARD members and those conducted by the CGIAR centres, of projects funded by EIARD members. The analysis in section 4 follows the structure set out in this table.

By reference to selected project evaluation documents, this analysis examines the outputs, outcomes and impacts of these assessments, including their influence on ARD decision making. The methodology is essentially a narrative review using the above analytical framework.

The long list of projects identified through the database searches and materials sent by EIARD members was screened to select those eligible for analysis. Impact assessments and evaluation reports were included if they explicitly considered outcomes and impacts – either in terms of the OECD-DAC definition or in the sense of attribution analysis.

It is important to note that in deciding which studies to analyse and which were ineligible, a number of single technology evaluations were excluded on the basis that these do not fit with the definition of ARD provided on page 9 of this document[58]. 178 projects were identified in Infosys+ responding to search terms 'impact' or 'evaluation' or the two combined. Most of these were reports on ARD projects testing or 'evaluating' technologies, but not the outcomes and impacts of the application of the technologies. However, 24 were impact assessments or evaluations with some

[58] that ARD should 'provide the necessary understanding of rural development situations, of the drivers and impediments for sustainable rural development, and the required knowledge and innovations for the development of smallholders' livelihoods'

Table 2 - Criteria for analysis of programme and project impact evaluations

Location and context
Is there a clear description of the type of agricultural research investment, its country location, the systems involved? Is there recognition of the complexity of the social, economic, policy and institutional context?
Scale and scope
Is the scale and scope of the evaluation clearly defined, e.g. whether based on a project or programme element versus evaluation of a research portfolio? Does it look beyond the outcomes arising from technology adoption to consider outcomes related to institutional development and partnerships?
Funding and commissioning relationship
Was the evaluation internally or externally commissioned? Was it directly funded by the EIARD member, or indirectly funded through CGIAR, or through intermediaries such as sub regional organisations? Was it jointly funded?
Purpose, objectives and type of evaluation/impact assessment
Are the *purposes* of the evaluation clearly defined? e.g. in relation to ARD investments to enhance their developmental impact; accountability to donors; to assess returns; to produce strategic lessons to inform future investment; to inform and influence policy and policy makers; to provide information for use in public awareness work etc. Does the evaluation take a holistic approach, considering outcomes for poverty, food security, natural resources use and conservation and sustainability? What *type* of evaluation is it; e.g. Outcome evaluation, Impact evaluation/ assessment?
Impact pathway
• Is there a statement and testing of an impact hypothesis? Is there a logical framework? • Is the programme theory or intervention logic mapped showing links between the intervention, its influence, the change promoted and impact targets, including definitions of beneficiary groups • Does the evaluation include a model or concept of innovation? •
Design, methods and tools
How is impact actually assessed or measured? e.g. Randomised experimental or quasi experimental design with counterfactual or control group; Before and after comparisons within an intervention group, with or without a baseline; time series. Are indicators defined and used for measurement? • Is the problem of selection bias addressed? e.g. through use of propensity score matching; pipeline approach, double difference; regression-based techniques, judgmental matching etc. • What kind of data collection methods are used - qualitative and/or quantitative methods? In what combination? E.g. interviews, focus groups, mapping, questionnaires, case studies etc. • Were participatory evaluation approaches used? Was there stakeholder consultation and participation? • What kind of data analysis – thematic narrative, modelling, systems analysis, cost benefit analysis? • Is there discussion of attribution and the limitations of the evaluation? Does it examine other factors, (social, political, economic and institutional) influencing the observed changes? • Is the evaluation design consistent with the purpose and objectives? •
Communication and dissemination
• Are there specific targeted recommendations to influence ARD decision making? • Is there critical review and comment on evaluation results by partners and beneficiaries? • Does the evaluation provide information for public awareness activities?

relevant information available. The DFID R4D database produced a further 11 projects not included in Infosys+, of which 9 were project or programme evaluations. Five projects were included from the 3ie list of impact assessments. Twenty evaluation reports were identified by EIARD members and nine were found in the Research into Use database. Out of the 224 projects identified, a total of 70 reports met the criteria for closer examination and analysis (see appendix 2). However, eliminating those which turned out not to be evaluations and those for which full reports were not available, left a total of just 50 projects. Further literature on evaluation approaches in ARD programmes was also consulted.

3 EIARD members policies and approaches to Impact Assessment of ARD

3.1 Principles and standards on evaluation and impact assessment

Among European donors there is a fair degree of consensus on the main purposes, principles and standards of evaluation[59] . The approaches to evaluation of development assistance of the European Commission (EC), GIZ, DANIDA, Sida and NORAD closely follow the OECD/DAC principles and standards. United Nations agencies follow the United Nations Evaluation Group standards[60] which identify 13 norms of evaluation, including independence, transparency, consultation and follow up, evaluation ethics and contribution to knowledge building. The Quality Standards for Evaluation specify criteria for the following areas;

- ***Impartiality and independence*** of the evaluation from policy processes and the delivery and management of development assistance.
- ***Credibility and transparency*** of process and reporting and sharing results.
- ***Usefulness*** – relevant findings and recommendations.
- ***Participation of funders and stakeholders***
- ***Donor cooperation*** – to avoid duplication of effort and to enhance joint learning and share information
- ***Evaluation programming*** –involving the users of the evaluation outputs.
- ***Design and implementation of evaluation***
- ***Reporting, dissemination and feedback***

However, these standards have been developed for evaluation of development assistance; there are no explicit criteria indicated for evaluation or impact assessment of ARD programmes as such, either at individual member level or for EIARD as a whole. However, it is assumed that similar norms would apply to ARD impact evaluations as to evaluation practice for development assistance more generally.

3.2 Impact assessment and country policies

In terms of approaches to impact assessment and methodologies used, it was difficult to find specific statements on these from EIARD members.

Two examples of countries policy on evaluation and/or impact assessment are set out below. Particular attention is drawn to the objectives and motivation behind the IA policies and the methodological and design recommendations which relate to these.

3.2.1 DFID UK

DFID's approach to impact evaluation[61] involves:

- Supporting and leading a shared international research collaboration with developing

[59] Quality Standards for Development Evaluation (OECD 2010).
[60] UNEG 2005, Norms for Evaluation in the UN system. New York: United Nations Evaluation Group
[61] DFID, 2009. Building the evidence to reduce poverty The UK's policy on evaluation for international development. Department for International Development (DFID) June 2009
http://www.dfid.gov.uk/Documents/publications1/evaluation/evaluation-policy.pdf
http://www.oecd.org/dataoecd/39/54/43183649.pdf

countries and partners such as the Gates Foundation and the World Bank to generate high quality impact evaluations that are relevant to and useful to developing countries;
- Founder member of the International Initiative on Impact Evaluation (3ie), set up to support rigorous impact evaluation of international development.
- Leading the Network of Networks on Impact Evaluation (NONIE) which has recently published authoritative guidance on impact evaluation approaches.
- Designing rigorous impact evaluations into the planning stage of many programmes that [DFID] funds or helps to fund.

DFID is the only EIARD member which provides a definition of impact evaluation in the sense of assessing attribution[62] ; '*Impact evaluation is a specialised type of evaluation which uses research methods to give us rigorous evidence on whether a policy, programme or project has actually changed people's lives and whether outcomes are directly attributable to the interventions'. For example, how do changes in nutrition and availability of health services delivered by a particular programme affect maternal mortality and health outcomes for young children in rural areas?*'

Evaluations are mainly concerned with the results and effectiveness of DFID's policies, programmes and partnerships –what, why and how results are achieved; how the programmes are working; who benefits/loses, effects on poverty (including intended and unintended effects); whether the policies and objectives are relevant to the ultimate aims of reducing poverty.

DFID have a 10 point implementation plan.

1. Strengthen the independence of central and decentralised evaluations
2. Meet Paris Declaration commitments to commission more evaluations jointly.
3. Improve evaluation of international partnerships. DFID supports use of partner systems for evaluation. Evaluations Department commissions evaluations of DFID's approaches to international partnerships with multilaterals and international NGOs.
4. Integrate the principle of policy coherence into our evaluations, including strengthening our evaluation links with other UK government departments.
5. Significantly increase number of decentralised evaluations in programmes and projects.
6. Support DFID staff to manage and develop their skills in evaluation and help to build M&E capacity in developing country governments and partners.
7. Establish clear mechanisms and incentives to use evaluations in decision making (indicators are that- annual evidence that DFID's key oversight committees are using relevant evaluation evidence in discussions on new plans and policies; that evaluation findings are feeding into and impacting on decision making and that there is increasing take up of training and support to build competencies in evaluation by point of mid term review.
8. Improve the quality of DFID Evaluations and strengthen quality assurance processes.
9. Increase staff skills in commissioning and management of high-quality evaluations.
10. Support more and better rigorous impact evaluation (through 3ie and programme of rigorous impact evaluations; publication of NONIE guidance on impact evaluation and delivery of rigorous impact evaluations of human development through DFID trust fund with World Bank and Spanish Government).

However, these plans are for DFID evaluations in general, not specific to ARD. The the majority of DFID funding for ARD is channelled through the CGIAR system, and this study identified few evaluations of directly funded ARD programmes carried out since this policy was formulated.

[62] DFID, 2009 op.cit. p 14

3.2.2 Netherlands

The leading principles of ARD in the Netherlands are to follow an agenda driven by demand from the South and to implement participatory approaches. The policy aim is to contribute to economic development, poverty reduction and sustainable use of natural resources in developing countries. The Ministry of Foreign Affairs supports overseas cooperation and international research, while the Ministry of Economic Affairs, Agriculture & Innovation (EL&I) funds Agricultural Research for Development (ARD) in the Netherlands. The majority of the research is carried out at Wageningen University and Research Centre (WUR) while applied ARD takes place at various applied research institutes. The Ministry of Foreign Affairs finances ARD at the International agricultural research centres[63] through the CGIAR by providing unrestricted core-funding to all 15 CG centres as well as 3 international education programmes in the field of agriculture.

The Policy and Operations Evaluation Department (IOB) of the Netherlands Ministry of Foreign Affairs is an independent department responsible for monitoring and evaluating all aspects of Netherlands' foreign policy and international cooperation. The purposes of evaluation are to enable accountability to parliament for policy and the allocation of resources, to derive lessons for the future by incorporating findings into the policy cycle and to target feedback to policy makers to improve policy design and implementation[64]. IOB has a staff of experienced evaluators and its own budget. When carrying out evaluations, it calls on the assistance of external experts with specialised knowledge of the topic under investigation. To monitor its own quality, it sets up a reference group for each evaluation, which includes external experts and interested parties from within the Ministry. Its approach has shifted from separate project evaluations to sector or thematic based evaluations. The reports are submitted to parliament and are in the public domain. In recent years, it has extended its partnerships with similar departments in other countries, for instance through joint evaluations. IOB also aims to expand its methodological repertoire, including a greater emphasis on statistical methods of impact evaluation. Among other topics, it has carried impact assessments on rural energy, water and sanitation.

The Netherlands will further this ever-changing agenda by strengthening links with other ARD actors through the Global Forum on Agricultural Research (GFAR) and regional fora; by building capacity and institutions at home and in the South; and by seeking more interactive means of upscaling that go beyond extension to the co-production of knowledge.

3.3 Processes and resource allocation in Impact Assessment

The questionnaire sent to EIARD country focal points asked for an estimate of their funding on impact assessment of ARD in the last five years. This proved very difficult to identify although it has been estimated that the overall EIARD donor investment in ARD in Sub Saharan Africa (SSA) amounted (in 2009) to US$163million per annum of which $106million per annum (65%) goes to the CGIAR[65]. It was difficult to separate funding for IA/evaluation from overall ARD funding although some members were able to estimate the proportion of their support to ARD

[63] Agricultural research for development is supported through core-funding to CGIAR institutes CIFOR, ICARDA, IPGRI, ICRAF, IWMI and the WorldFish Centre.
[64] Dutch Ministry of Foreign Affairs, Policy and Operations Evaluation Department
http://www.minbuza.nl/en/ministry/policy-and-budget/evaluation-of-foreign-policy-spending/policy-and-operations-evaluation-department-iob.html

which is allocated to impact evaluation. For example, Germany estimated their overall funding at about 15 million euro per year of which approximately 1-2% would go into impact assessment of large projects and programmes as well as small grants. An additional 5 million euro per year is for support of approximately 40 scientists seconded to the CGIAR centres and 1-2% of this amount is for their performance assessment with partner institutions. Switzerland estimated that 16.5% of their total ARD investment was allocated for evaluation and impact assessment.

The EIARD member countries have different relationships to evaluation and impact assessments, depending on whether the majority of their ARD funding is provided through direct support to programmes and projects or through multilateral organisations, such as the CGIAR (see figure 3). Some countries combine the two strategies. Within the CGIAR funding category there is a distinction between restricted and unrestricted funding. Those providing unrestricted funding to the CGIAR as well as some of those giving restricted funds, rely solely on the impact assessment procedures within the CGIAR system, while others combine the system level reporting with their own evaluations.

The pattern of funding of EIARD members shown in Figure 3, includes those countries with little or no ARD (Box 1); those which almost exclusively fund ARD within the CGIAR centres or through other bodies such as ASARECA and FARA[66] (Box 2); those that combine a majority of CGIAR funding with their own ARD investments (Box 3), and those for whom the majority of ARD is directly funded (Box 4). There was a high level of support to the CGIAR centres, both core and restricted funding (15 out of 17 responses). The largest donors to the CGIAR are the EC and the UK, followed by Germany, Switzerland and Norway. In addition, 13 of the 17 EIARD respondents fund their own ARD programmes directly.

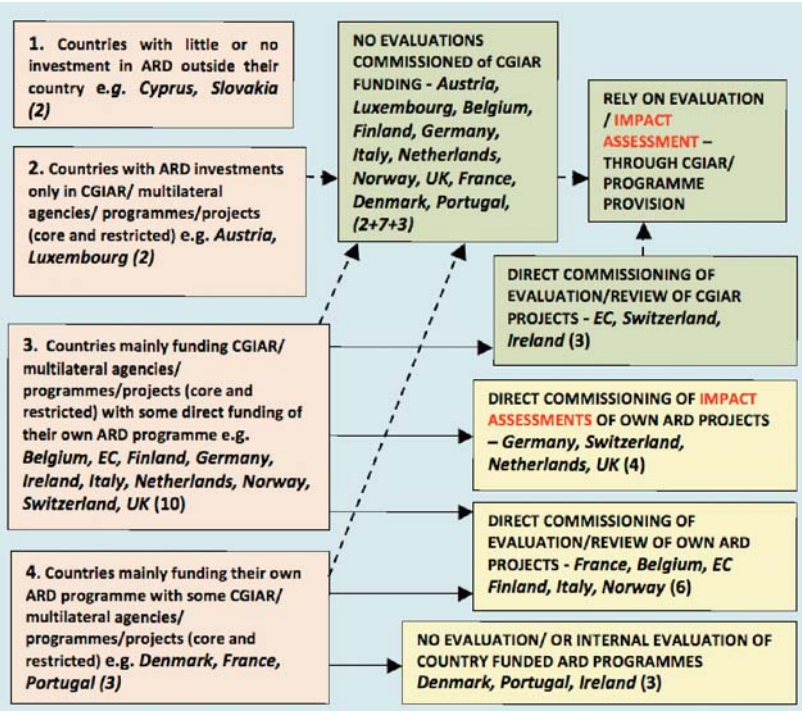

Figure 3 - Different EIARD member positions on funding ARD and evaluations/ impact assessments (information for 17 members

[65]James Morton & Co. (2010, Analysis of donor support to CAADP Pillar 4. Draft final report to the European Initiative for Agricultural Research for Development (EIARD)
[66] Association for Strengthening Agricultural Research in East and Central Africa - ASARECA. Forum for Agricultural Research in Africa -FARA

The responses revealed different levels of engagement in evaluation and impact assessment for CGIAR funding and country funded ARD programmes. For example, for CGIAR funding, the EC has evaluated its contribution[67] and examined the impact of CGIAR outputs[68] and Switzerland has conducted impact assessments. Switzerland has also conducted evaluations and reviews and Ireland has recently commissioned a review of its overall support to CGIAR, although in terms of specific projects or Centres it has relied on the SPIA impact assessments rather than duplicating by carrying out its own evaluations. The rest of the EIARD members funding the CGIAR rely exclusively or mainly on the evaluation arrangements within the system, typically requiring only annual reports from the funded programme or centre.

Among those with their own ARD programmes, Germany, Switzerland, Netherlands and the UK have carried out impact assessments, while Belgium, Finland, France, Italy and Norway have conducted evaluations or reviews of their own ARD projects.

[67] Ooijen, Rudy, and David Coombs (2007) Evaluation of EC Contribution to the Consultative Group on International Agricultural Research (CGIAR), Final Report (Volume I) ECORYS Nederland BV
[68] EC 2011. Practical Application of CGIAR Research Results by Smallholder Farmers.

4 Analysis of selected studies of Impact Assessments

This section presents the findings from an analysis of the studies gathered for this study. This does not constitute the scope of a full systematic review, but rather draws findings in a qualitative manner from a fragmented and somewhat elusive evidence base. We consider the location and context, scale and scope, funding and commissioning relationships, the purpose, objectives and types of impact assessment/evaluation; impact pathways; design, methods and tools; and communication and dissemination.

4.1 Location and context

Given the broad definition of ARD a key question is whether the evaluation studies contain a clear description of the type of agricultural research investment, its country context, including its agro ecological and policy and institutional context. A large number of the studies initially identified had a narrow focus on technology adoption rather than the outcomes for farmers and other beneficiary groups in particular situations. Best evaluation practice involves contextualising the findings in order to learn what works in what situations.

4.2 Scale and scope

From the studies gathered by the research team slightly more than half were single project evaluations or impact studies, or focused on one country, whereas the rest were larger programmatic evaluations, involving multiple stakeholders and covering multiple countries, while a small number were more thematic or synthetic in scope.

Of the studies selected for further analysis there were 23 single project or single country studies; 21 programmatic reviews and 6 syntheses or thematic reviews.

4.3 Funding and commissioning relationship

The evaluation standards place considerable emphasis on the impartiality and independence of the evaluation. It should be separate from policy processes and the delivery and management of development assistance. This can be achieved by contracting an external body or consultant to design and conduct an impact evaluation, or having a separate structure within the organisation which is responsible for evaluation.

Within the evaluation reports consulted, the commissioning relationship was rarely specified and the EIARD member countries focal points did not provide much information in response to the question on who conducts impact evaluations.

4.4 Purpose, objectives and type of evaluation/impact assessment

The evaluation reports analysed covered a range of objectives; some were looking specifically at developmental impact and accountability, some were seeking to use information for strategic planning and others were more concerned with learning. Evaluations varied greatly in the emphasis placed on measurement and attribution of change arising from ARD. For example, the SDC funded 5 Year Ex-Post Impact Study on the POSTCOSECHA Programme Central America had as its

Table 3 - Scale and scope of the selected studies

Single project, focused theme and/or country evaluation Number of projects 23	Programmatic evaluations, with multiple components and countries: No. Progs. 21	Meta- evaluations, syntheses. Number of reports 6
Impact of irrigation development on rural poverty and the environment, Ethiopia (1)	Sub-Saharan Africa Challenge Programme project, involving multiple stakeholders and 9 learning sites in Southern, East and West Africa (2)	Community animal health services for improving household wealth and health status of low income farmers (9)
Participatory monitoring and output assessment of rural regeneration and sustainable agriculture in Brazil (focusing on Project Paraiba, IIED) (4)	Review of rice wheat consortium for the IG plains (7) / Rice impact study project, Indo-Gangetic Plain (CABI) (3)	Meta-analysis of evaluations conducted by Directorate General for International Cooperation and Development (25)
Using livestock to improve livelihoods of landless and refugee affected livestock keepers in Bangladesh and Nepal (DFID livestock production programme) (6)	The National Agriculture and Livestock Extension Programme (NALEP) Programme Impact Assessment (13)	Monitoring and Evaluation: Pathways for Change - A Summary of Monitoring & Evaluation Experience from the Renewable Natural Resources Research Strategy (52)
Evaluation of striga, stemborer and soil fertility management techniques in Kenya, Tanzania, Uganda (8)	Evaluation of EC contribution to CGIAR 2007 (19)	Synthesis of Monitoring and Evaluation Experience in the Renewable Natural Resources Research Strategy (RNRRS) (54)
Tomato leaf curl virus disease resistant tomatoes (10)	Practical application of CGIAR research results by smallholder farmers (20)	Results-Based Approach in Finnish Development Cooperation (58)
Integrated Economic and Social Analysis to Access the Impact of Vegetable and Fishpond Technologies on Poverty in Rural Bangladesh (11)	Evaluation of Belgian aid to Rwanda (21)	The Sustainability Dimension in Addressing Poverty Reduction: Synthesis of Evaluations (59)
Testing the Impact of Farmer Field Schools, using survey data from Peru (12)	Evaluation of direct aid from Belgium – 4 case study countries (22)	
Assessing upland rice varieties for India (DFID RNRRS research project) (15)	Joint evaluation of EU AID from Belgium, France, Denmark and Luxembourg (23)	
Rainfed Rabi Cropping in Rice-fallows of Nepal, DFID RNRRS (16)	A cheval donne – on ne regarde pas les dents . Les mechanisms et les impacts de l'aide vus par les practiciens Nigeriens (24)	
Institutional innovations in India's crop improvement system (DFID Plant Sciences Research Programme) (17)	Ex post impact study of the POSTCOSECHA programme, Central America (26)	
From Research into Use: The innovation trajectory of sleeping sickness control in Uganda: research knowledge in its context. Discussion Paper 09 (18) / Monitoring & Evaluation of a Public-Private Partnership Stamp out Sleeping Sickness Case Study (56)	Research to impact; case studies for NRM for irrigated rice in Asia (27)	
Water saving in rice production - Study of dissemination, adoption, and early impacts of Alternate Wetting and Drying Technology in irrigated rice in Bangladesh (28)	Evaluation of Swiss Agency for Development and Cooperation (SDC) research related activities (29)	
Impact assessment of push-pull technology developed and promoted by Icipe and partners in East Africa (31)	Evaluation of SDC's contribution towards Biodiversity: Impact in the Andean Region (30)	
Socio-cultural factors influencing the adoption of ecologically based rodent pest management in Vietnam (33)	Study of the effectiveness of SDC cooperation in the agricultural sector (32)	
Improvement of rice productivity in South and Southeast Sulawesi (34)	Assessing the farm level impact of genetically improved	

Table 3 - continued

Single project, focused theme and/or country evaluation Number of projects 23	Programmatic evaluations, with multiple components and countries: No. Progs. 21	Meta- evaluations, syntheses. Number of reports 6
Balanced nutrient management systems in West Africa, focusing on northern Guinea savanna of Nigeria (35) Assessing the impact of banana bacterial wilt, Uganda (37) Impact of production and marketing of freshwater aquatic products on rural livelihoods. (44) The impact of organic cotton cultivation on the livelihood of Indian smallholders (48) Monitoring, impact assessment and learning for research into use: Nigeria (66) The Impact of Improved Maize Varieties on Poverty in Mexico: A Propensity Score-matching Approach (68) Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands (69) / Conservation Agriculture in Africa - CIRAD (14) Technology adoption, productivity and specialisation for Uruguayan breeders: evidence from an impact evaluation (70)	farmed tilapia (36) Determination of high-potential aquaculture development areas and impact in Africa and Asia (38) Impact of agricultural sector market reforms on smallholder farmers in Benin and Malawi (42) Impact of aquatic animal health strategies on Livelihoods of poor people in Asia. (45) Understanding rural livelihoods systems for rice research prioritization and impact assessment' (49) Assessing the contribution of diversified Musa genetic resources to poverty reduction, environmental sustainability and gender equality in rural communities (51) Evaluation of DFID RNRRS (55)	

objectives: i) to provide facts and explanations for accountability purposes and ii) to identify key information for organizational learning in order to promote the approach elsewhere[69].

The types of evaluation also ranged from those reporting against indicators and targets, and taking a broad 'plausible' linkages approach to examine the case for attributing change to the research intervention. Others were designed with a more rigorous experimental approach, with the use of a counterfactual and efforts to measure the extent of impact. Of the 44 project or programme studies, just six were more formal impact evaluations.

4.5 Specification of impact

4.5.1 Impact pathways, logical frameworks and indicators

Many donor funded projects use logical frameworks in planning, but not all evaluations indicate the planned outcomes of the research or state the hypothetical impact pathway for the research. There is a distinction between the use of the term 'impact pathway' as an exercise to develop a practical strategy for delivery and dissemination, compared to the usage of the term which seeks to map out the *logical links and hypothetical* connections between the research activities and their impacts, including contextual factors and influences that may shape outcomes and impacts.

Among the sample of projects analysed it was unclear how many projects had logical frameworks as the basis of project design and against which the project could be evaluated. Very few of the single project/country project studies tease out an impact hypothesis or clearly map the logic linking the intervention, its influence and the change promoted.

The findings of an evaluation[70] of evaluations of Swiss Agency for Development and Cooperation (SDC) research projects came to similar conclusions. Firstly, the analyses mainly concentrated at output rather than outcome level; only a third of the projects had logical frameworks and only two had an explicit model of change. '*Generally individual projects do not report on other parts of the innovation system, and do not report on what other inputs are necessary for their work to achieve an impact.*'

Only a few programme/multi-country evaluations (e.g. SSA Challenge Programme, assessment of the farm level impact of genetically improved farmed tilapia and the report on the Effectiveness of Swiss development cooperation in the agricultural sector in 2010[71]) specified impact pathways

[69] Martin Fischler 2011. 5 Year Ex-Post Impact Study on the POSTCOSECHA Programme Central America. Intercooperation (26)
http://www.postcosecha.net/en/Home/Documentation/Publications/Publications_in_English

[70] Swiss Agency for Development and Cooperation 2010. Evaluation of SDC's Research Related Activities, Bern, March 2010 https://ext.d-nsbp-p.admin.ch/NSBExterneStudien/19/attachment/fr/70.pdf (29)

[71] Uzo Mokwunye and Jim Ellis-Jones 2010, SSA Challenge Programme, Internal Review Report FARA (2)
http://www.dfid.gov.uk/r4d/projectsAndProgrammesResults2.asp?search=related%20Documents&Outputs=Yes&projectID=60686.
Asian Development Bank 2005. An impact evaluation of the development of genetically improved farmed tilapia and their dissemination in selected countries. Operations Department (36). http://www.extension.org/mediawiki/files/f/f5/An_Impact_Evaluation_of_the_development_of_Genetically_Imprv.pdf
Swiss Agency for Development and Cooperation n/d Report on the Effectiveness of Swiss development cooperation in the agricultural sector in 2010 Technical analyses (32)
http://www.deza.admin.ch/en/Home/Effectiveness/Reports_on_effectiveness_and_annual_reports

or results chains with clear linkages. In the sample there are studies which do not explicitly state their method or specify the links between research inputs, outputs, outcomes and impacts.

The study '*Determination of high potential aquaculture development areas and impact in Africa and Asia*'[72] is a stakeholder evaluation of a project at the World Fish Centre, funded by Germany (GTZ). The WorldFish project outputs are identified: i) an integrated knowledge base of freshwater pond aquaculture systems and practices, and driving factors for their adoption and continued development; and ii) an analysis package that can be used in decision support to identify places and situations for which freshwater pond aquaculture is feasible, and the nature of constraints requiring interventions to realize the potential of the target areas. A GTZ representative suggested an impact assessment during the mid-term review of the project. While too early for a classical ex-post impact assessment including measurement of impacts on ultimate beneficiaries, an evaluation of project performance, and particularly outputs, but also the likelihood of achieving expected outcomes (e.g. changes in awareness, attitude and knowledge) based on national stakeholder views was thought to be of use. This is a hybrid M&E document and stakeholder evaluation that looks forward to likely uptake and some outcomes. Figure 4 reproduced below shows the impact pathway for this research. It is fairly simple and straightforward.

Figure 4 - Example of an impact pathway – Aquaculture development

A more complex diagram is provided by the evaluation of the Stamp Out Sleeping Sickness control project in northern Uganda, funded by DFID[73]. A preliminary working model together with questions arising was developed to summarise how uptake and impact might be achieved and to guide areas of investigation during Case Study development, see Figure 5 from annex 3 of the report, reproduced below.

The third example is from a study of dissemination, adoption, and early impacts of Alternate Wetting and Drying Technology in irrigated rice in Bangladesh co funded by IRRI and the Advisory Service on Agricultural Research for Development (BEAF) of GTZ. This was an IRRI project. The

[72] Diemuth E. Pemsl and Terence Too, June 2008. Determination of high-potential aquaculture development areas and impact in Africa and Asia (38)
[73] . John Morton, 2010. The innovation trajectory of sleeping sickness control in Uganda. Research knowledge in its context. Discussion Paper 09. Research into Use Programme (18)
http://www.researchintouse.com/resources/riu10discuss08ssickcntrl-ug.pdf

paper[74] describes and analyses the entire process of introducing the innovation in Bangladesh, including how Bangladeshi institutions and organizations have become involved in the process of promoting the new technology and how it has been up- and out-scaled. The study looked at short term impacts since the dissemination and adoption of AWD are still very recent in Bangladesh and hence aggregated impacts could not yet be measured. The assessment includes economic, social and cultural impacts on the adopter level. Figure 6 below shows the research to impact pathway approach which is promoted by IRRI. In this study they provide a whole section on the Research to Impact Pathway Concept. Using the 'impact pathway' they illustrate the way research contributes to agricultural change and broader social, economic and cultural impacts.

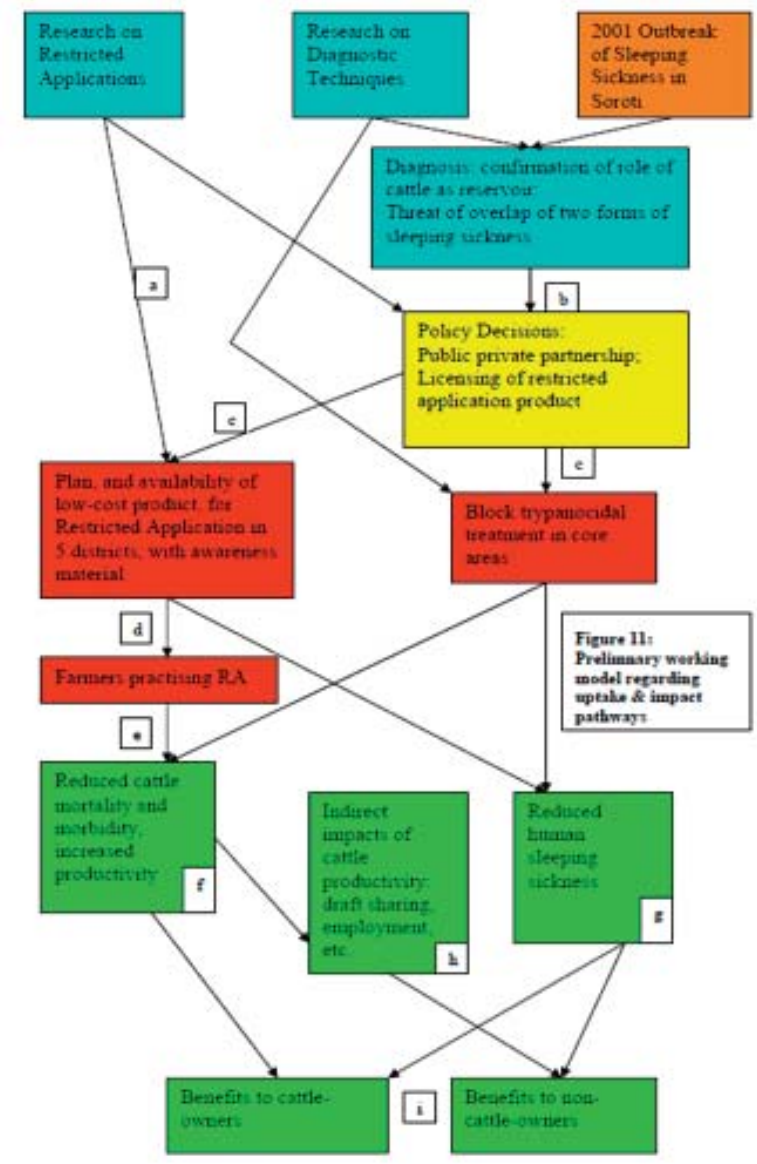

Figure 5 - Model of uptake and impact pathways - Sleeping sickness control project. Source John Morton, 2010

[74] Dr. Ekkehard Kürschner, Christian Henschel, Tina Hildebrandt, Ema Jülich, Martin Leineweber, Caroline Paul (2010) Dissemination, Adoption and Short Term Impacts of Alternate Wetting and Drying (AWD) in Bangladesh. Dhaka/Berlin (28). http://edoc.hu-berlin.de/series/sle/241/PDF/241.pdf

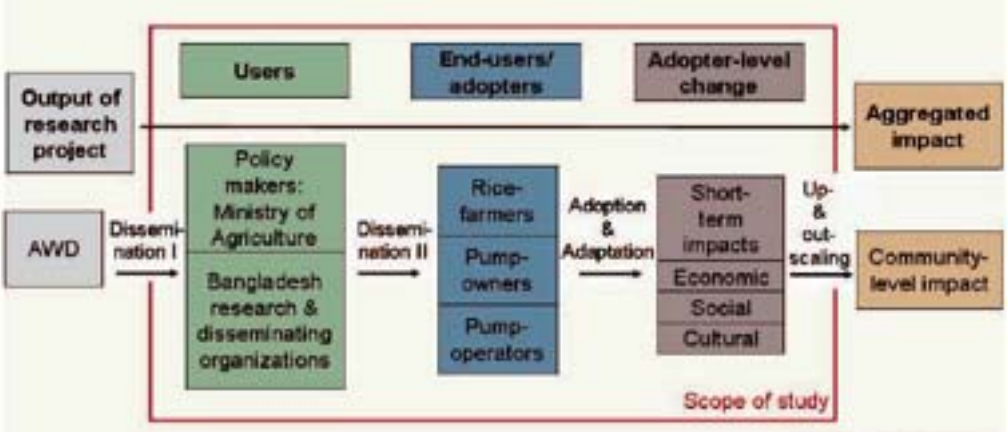

Figure 6 - Promotion of alternate wetting and drying technology in Bangladesh - research to impact pathway

From the overall analysis it is clear that impact pathway analysis has not been a widely used tool. It was utilised for several years as part of the CGIAR Medium Term Plan requirement but functioned more in the sense of uptake pathways for dissemination rather than a depiction of the causal pathways, relationships, influences through which change is brought about. The use of impact pathways and diagrammatic representations of theories of change is likely to increase as more emphasis is being placed on the use of this approach in impact evaluation and the need for tools which complement project logical frameworks is recognised.

Table 4 Number of evaluations using impact pathways, frameworks and indicators

Total projects/ programmes	Evaluations using impact pathways	Evaluations using logical framework and/or indicators	No explicit usage of impact pathways or logical frameworks
44	6	5	33

The sample of 44 project or programme reports (not the thematic reviews) were also examined more generally in terms of their use of logical frameworks and impact indicators. Table 4 shows the overall distribution of the reports in terms of their use or otherwise of impact pathways/theory of change, logical frameworks and indicators.

There were five examples of evaluations which specified the use of logical frameworks and/or indicators as tools in evaluation. The evaluation of Community-based Irrigation Management in Ethiopia: Strategies to Enhance Human Health, Livestock and Crop Production, and Natural Resource Management[75] was carried out by Mekelle University, ILRI, IWMI and Austrian institutions, with funding contributed by Austrian Development Agency. The analysis and detailed case studies focused on the causal linkages among irrigation investment, poverty reduction, food security, economic development, environmental externalities, and health under Ethiopian conditions. It used comparative indicators to evaluate the performance of three irrigation schemes in Tekeze basin.

[75] Behailu, M, Abdulkadir M, Mezgebu A and Yasin, M 2005 Community based Irrigation Management in the Tekeze Basin: Perfomance Evaluation: A Case study on three small-scale irrigation schemes (micro dams) (1) http://www.iwmi.cgiar.org/assessment/files/word/ProjectDocuments/ILRI/Atinkut_Community%20Based%20Irrigation%20_ProjectReport.pdf

The study, *Reaping the Benefits: Assessing the impact and facilitating the uptake of resource-conserving technologies in the rice-wheat systems of the Indo-Gangetic Plain*[76], carried out a livelihood impact assessment of improved technologies on social well-being, system productivity and sustainability at selected benchmark sites. The project was a collaboration between CABI, CIMMYT and national research organisations and NGOs in Nepal, Pakistan, India and Bangladesh, funded by DFID. The assessment is structured closely around the goal, purpose and outputs of the logical framework.

The Participatory monitoring and output assessment of rural regeneration and sustainable agriculture in Brazil[77] used a participatory learning approach to monitoring and impact assessment. The study was intended to improve M&E approaches for application to RNR research programmes and projects. It sets out clear ideas on the potential and pitfalls of a more open-ended (indicator-less) approach, and one that focuses on specific indicators. In particular, the development and application of a standard set of indicators, both qualitative and quantitative, was difficult to achieve due to the high level of diversity of environmental and organisational characteristics.

An example of an evaluation study where specific outcome indicators have been used is *Technology adoption, productivity and specialisation for Uruguayan breeders: evidence from an impact evaluation*[78]. This aimed to identify the average impact of a Livestock Pilot Project on the intended development outcomes as they were defined in the original design of the project. In particular, the project design identified two specific outcomes indicators for component 1: (i) the Reproductive Efficiency Index (REI and (ii) the rate of adoption of managerial practices.

4.5.2 Disaggregation of impacts by beneficiary group

The evaluation reports were also examined for the extent to which they disaggregated the impacts according to different beneficiary groups. As shown in table 5, 33% of the evaluations had some indication of disaggregation of beneficiaries and attempted to assess how different levels of benefits were experienced by different members of the community. Disaggregation was mainly in terms of different wealth or asset status; fewer reports discussed gender aspects of impact in any detail. This is consistent with the findings of another study for EIARD, on making ARD more pro-poor[79].

[76] Proceedings of Technical Review and Planning Workshop: "Assessing the Impact of Resource Conserving Technologies in the Indo-Gangetic Plain Identifying Agricultural Knowledge Systems and Overcoming Blockages to Enhance Uptake of Agricultural Technologies to Optimise Pro-Poor Development"; 5-7 May 2004, Hotel Eastern Residence, Banani, Dhaka, Bangladesh. . Report Compiled and edited by Sam L J Page and Tahseen Jafry (3) http://www.dfid.gov.uk/r4d/PDF/Outputs/Misc_Crop/Dhaka_workshop_report.pdf

[77] Sidersky P and Guijt I. 2000. Experimenting with participatory monitoring in north-eastern Brazil: The case of AS-PTA's Projeto Paraiba. In: Estrella et al. (editors). Learning from Change: Issues and experience in participatory monitoring and evaluation. London: IT Publications (4) http://www.dfid.gov.uk/r4d/searchresearchdatabase.asp?ProjectID=1341

[78] Fernando Lopez and Alessandro Maffioli 2008. Technology adoption, productivity and specialisation for Uruguayan breeders: evidence from an impact evaluation. (70) http://www.3ieimpact.org/page.php?pg=details_impact&id=149

[79] Pound, B, Michiel van Dijk, Yuca Waarts and Essie Apenteng 2011, Making ARD more pro-poor; Improving accessibility and relevance of results to the poorest. Agrinatura.

[80] Proceedings of Technical Review and Planning Workshop: "Assessing the Impact of Resource Conserving Technologies in the Indo-Gangetic Plain; 5-7 May 2004, Hotel Eastern Residence, Banani, Dhaka, Bangladesh. (3) http://www.dfid.gov.uk/r4d/SearchResearchDatabase.asp?OutputID=173983.

Olaf Erenstein 2009, Zero Tillage in the Rice-Wheat Systems of the Indo-Gangetic Plains A Review of Impacts and Sustainability Implications IFPRI Discussion Paper 00916 November 2009 (7) http://www.ifpri.org/sites/default/files/publications/ifpridp00916.pdf

Table 5 Project and programme evaluations disaggregating beneficiaries.

Total projects/ programmes	Evaluations with disaggregation of beneficiaries – wealth status and gender	Evaluations with no apparent disaggregation of beneficiaries
44	15	29

The Rice impact study project, Indo-Gangetic Plain (CABI)[80] conducted a livelihood impact assessment of resource conserving technologies, with farmers in villages in four countries across the Indo-Gangetic Plain: India, Pakistan, Nepal and Bangladesh. It examined the effect of the technologies on social well-being and system productivity for each social group within each community at selected benchmark sites. Particular attention was paid to marginal and landless farmers, including women. The study showed that traditional technology dissemination methods discriminate against risk-averse marginal and landless farmers, whereas optimal uptake can result from dissemination methods that focus on the needs of the poorest.

The project 'Using livestock to improve livelihoods of landless and refugee affected livestock keepers in Bangladesh and Nepal[81]' similarly explored differences between landed and landless people and their aspirations and constraints, including those of female headed households. The report 'The Impact of Organic Cotton Farming on the Livelihoods of Smallholders' on a project in central India examines indicators (education, caste, housing, land, equipment, off-farm income and irrigation, to explore differences between organic and non organic farmers[82].

Gender disaggregation of findings was found in a number of other projects (e.g. Evaluation of striga, stemborer and soil fertility management techniques in Kenya, Tanzania, Uganda[83]; Integrated Economic and Social Analysis to Assess the Impact of Vegetable and Fishpond Technologies on Poverty in Rural Bangladesh[84]; 5 Year Ex-Post Impact Study on the POSTCOSECHA Programme Central America[85]).

Other forms of differentiation include differentiating market actors, specifically types of traders, their specialisation and concentration (e.g. Impact of agricultural market reforms on smallholder farmers in Benin and Malawi[86]). Similarly, the report on Impact of production and marketing of freshwater aquatic products on rural livelihoods[87] distinguishes between poorer fish farmers, poorer non-fish farmers, richer fish farmers, richer non- fish farmers in each of 12 villages.

[81] Planning and evaluation with landless people. http://www.dfid.gov.uk/r4d/PDF/outputs/R8109b.pdf DFID Livestock Production Research Project (6)
[82] Eyhorn, Frank, Paul Mäder, Mahesh Ramakrishnan 2005, The Impact of Organic Cotton Farming on the Livelihoods of Smallholders Evidence from the Maikaal bioRe project in central India (48) http://www.fibl.org/fileadmin/documents/en/development-cooperation/production-systems/executive_summary.pdf
[83] Z. R. Khan 2005. Promotion and dissemination of Integrated Pest and Soil Fertility Management Strategies to combat striga, stemborers and declining soil fertility in the Lake Victoria basin R8449 Final Technical Report. (8) http://www.dfid.gov.uk/r4d/PDF/Outputs/CropProtection/R8449_FTR.pdf
[84] Kumar, Neha and Agnes R. Quisumbing 2010, Access, Adoption, and Diffusion; Understanding the Long-term Impacts of Improved Vegetable and Fish Technologies in Bangladesh, Poverty, health and nutrition division IFPRI Discussion Paper 00995 June 2010 (11) http://www.3ieimpact.org/admin/impact_evaluations/Access,%20Adoption,%20and%20Diffusion.pdf
[85] Fischler, 2011. (26)
[86] Gabre-Madhin, Eleni, Marcel Fafchamps, Richard Kachule, Bio Goura Soule, Zahia Kahn 2001. Impact of agricultural market reforms on smallholder farmers in Benin and Malawi. Final Report Volume 2. Submitted to the Deutsche Gesellschaft Für Technische Zussammenarbeit (GTZ) (42) http://www.ifpri.org/sites/default/files/publications/v2.pdf
[87] Golam Faruque, 2007 An exploration of impacts of aquaculture production and marketing on rural livelihoods in three regions in Bangladesh. Thesis submitted for the degree of Doctor of Philosophy. University of Stirling. (44) https://dspace.stir.ac.uk/bitstream/1893/253/1/PhDThesis_GolamFaruque.pdf

Among the broader programme evaluations there is also consideration of beneficiary groups. In the Evaluation of EC contribution to CGIAR 2007[88], the wide range of actual and potential beneficiaries from the projects is noted. Depending on the type of project, these are ultimately the small-scale farmers and consumers, however, other beneficiaries more immediate beneficiaries are the international and regional organisations, Universities and other research bodies, NARS and extension services NGOs, traders, commercial seed farmers and farmer groups.

The focus on beneficiaries may be stronger where there is a more developmental rather than research focus. The impact assessment of the National Agriculture and Livestock Extension Programme (NALEP) Phase I in Kenya (2006)[89] commissioned by Sida, examines the impacts on people living with HIV/AIDS; the gender dimensions of impact and the extent of inclusion of the poor and vulnerable. The report notes the constraints faced by field extensionists in finding the time for visits to individual affected households (necessary because of public stigma) and recommends further training and linkages with other organisations active in the field. The evaluation examined the levels of women's participation in different activities and in decision making through a number of indicators on gender and participation. It noted that the project had defined categories of the poor and vulnerable and then involved them in the project and monitored their progress over time.

4.5.3 Evaluations and the concept of innovation

The projects reviewed, in so far as they articulate a concept of innovation, have rather different definitions of this term. For example, the IRRI study of Alternate Wetting and Drying Technology in Bangladesh follows the definition by Rogers[90]. He sees innovation as an '*idea, practice or object that is perceived as new by an individual or other unit of adoption*'. The actual diffusion of an innovation refers to the '*process by which an innovation is communicated through certain channels over time among members of a social system*'[91]. This process can be actively promoted or disseminated or be spontaneous, for example, through farmer-to-farmer diffusion.

This contrasts with more recent definitions of innovation and innovation systems which see innovation itself as a process rather than a specific invention, engaged in by multiple stakeholders playing different roles in bringing about technological change, together with the institutions necessary to support it. Hall *et al*[92] argue that impact assessment research which is heavily reliant on measurement of economic impact, fails to provide research managers with critical institutional lessons concerning ways of improving research and innovation as a process. The innovation systems analytical framework provides an alternative approach to institutional learning and looks closely at the actors, partnerships, skills and interactions in particular institutional contexts – rather than examining changes in technology user groups alone.

[88] Ooijen, Rudy, and David Coombs 2007 Evaluation of EC Contribution to the Consultative Group on International Agricultural Research (CGIAR), Final Report (Volume I) ECORYS Nederland BV (19)
[89] Melinda Cuellar, Hans Hedlund, Jeremy Mbai, Jane Mwangi 2006. The National Agriculture and Livestock Extension Programme (NALEP) Phase I Impact Assessment. Sida Evaluation 06/31. (13)
http://www.sida.se/Documents/Import/pdf/0631-The-National-Agriculture-and-Livestock-Extension-Programme-NALEP-Phase-I-Impact-Assessment.pdf
[90] E Rogers, 2003 p12
[91] Ibid p5
[92] Hall, A.J., Rasheed Sulaiman, V., Clark, N.G., 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 (2003) 213–241

Only two evaluation reports discussed innovation in the sense outlined above; the Stamp out Sleeping Sickness report and Institutional innovations in India's crop improvement system.[93] The latter defines an institutional innovation as a change in the institutions (rules of the game) through which a new technology is developed and becomes adopted by end-users.

The majority of the evaluations consulted made little reference to innovation. Several discussed innovation in the sense of a new technology. The Sub Saharan Africa Challenge Programme internal review report[94] refers to innovation platforms (IPs). These appear to be conceptualised as localised innovation groupings around households and communities. The project has 'intervention' village IPs and counterfactual comparison villages and households and has developed process and impact indicators.

The recommendation of the EIARD task force to include a model or concept of innovation in evaluations, appears to have far to go and there remain important requirements for understanding and learning about the institutional context of agricultural research and development processes.[95]

4.6 Design, methods and tools

The type of evaluation, evaluation design and methods were analysed for the projects and programmes. The task was to examine how change resulting from the project was actually assessed.

The majority of the evaluation reports consulted were actually *outcome evaluations* in that they attempted to trace the outcomes of the project interventions, but not to measure or attribute *impact*. The distinction was made (pages 7 and 8 above) between evaluating outcomes and evaluating impacts: the former draws evidence from monitoring data and systematic stakeholder feedback evidence and the latter (evaluating impacts) incorporates a counterfactual and therefore allow specific comparisons to isolate the effect of the intervention and hence attribution of impact.

The studies were grouped into three types: those reporting simple outcome evaluations largely focussing on beneficiary groups (38.64%); an intermediate group of outcome evaluations (47.73%) which made some comparisons e.g. 'before' and 'after' comparisons or compared a group of project participants with a similar group of non participants. Amongst these outcome evaluations were a small number of highly participatory studies. The third and smallest group (13.64%) were impact studies, characterised by their degree of rigour and emphasis on measurement of change, incorporating counterfactuals and using sophisticated statistical techniques to overcome sampling biases.

4.6.1 Simple Outcome evaluations

The simplest form of evaluation in this group typically involves a survey of beneficiaries to explore the changes brought about by their engagement with the project. The investigation may

[93] Morton, 2010 (18). Czech Conroy, 2009 Institutional innovations in India's crop improvement system: Rainfed Agriculture Impact Assessment Study No. 5 September 2009 (16)
http://www.researchintouse.com/resources/riu09impact4nepal-rice.pdf
[94] Mokwunye and Ellis-Jones, 2010. (2)
[95] Learning from the Renewable Natural Resources Research Strategy Pathways for change: monitoring and evaluation
http://www.dfid.gov.uk/r4d/pdf/ThematicSummaries/Brief3_Pathways_for_change_monitoring%20and%20evaluation.pdf

Table 6 - Evaluation methodologies of reviewed projects and programmes.

Methodology of evaluation	No. of Project evaluations	No. of Programme/ multi country evaluations
Simple outcome evaluation 38.6%	9	8
Outcome evaluation with 'before' / 'after' or 'with/without/ comparisons 47.7%	9	12
Impact evaluations (quasi experimental design, with counterfactual) 13.6%	5	1

range from issues of technology adoption and performance, to exploring changes in incomes, human health, the environment etc. Differential impacts on women and the poor may also be considered. For projects aimed at institutional change, these dimensions were the main focus. Methodologies used were semi structured interviews with key informant, questionnaire surveys and participatory assessment methods (alternatively or in combination) and review of secondary data and documents. Participatory methods included focus group discussions, development of participatory evaluation criteria, ranking and scoring of technology performance, 'peer review' evaluation by exchange visits (farmer to farmer and research/extension to research extension). Reliability was sought by 'triangulating' different sources of both quantitative and qualitative, different stakeholders consulted and key documents. In several cases, villages and participants for interviews were selected purposively according to agreed criteria, such as location and distances to service and market opportunities, duration of interventions, accessibility etc.etc.

The methodology used by *programme* outcome evaluations was a little more complicated since it had to take into account multi locational interventions and multiple partners, or a more complex set of interventions in a single location. In addition to the methods above, qualitative case studies (see Box 2), electronic questionnaire surveys and stakeholder evaluations were used.

Some of these evaluations – both project and programme - used an impact pathway or results chain as a framework against which to assess the processes and outcomes[96].

This type of evaluation is often intended to contribute to project learning and future performance, to provide accessible information for local stakeholders or in preparation for a subsequent phase. Its value is significantly enhanced when there is good monitoring data providing evidence for the trajectory of change over time.

4.6.2 Outcome evaluation with some comparisons

This group of project and programme evaluations were designed to provide information on the differences made by the project, through structured comparisons. The main types of comparison are 'before' and 'after' the project intervention - an approach which requires some form of baseline data collection which can be compared with a final survey; a 'with' and 'without' project intervention which compares participating groups with similar non participating ones; and a sequential or 'pipeline' design which compares participants with a similar group or groups which will participate *in the future*. When 'before' and 'after' and 'with' and 'without' are combined, this

[96] Projects with impact pathways were numbers 18, 32, 36 and 38 in table 1.

is known as '*double difference*' as it allows for comparison of the participating group and non participating group, assuming that the groups remain stable in composition and that external influences are the same for both during the period of intervention. Three projects[97] in this category used the before and after framework and conducted baseline and post-baseline surveys.

Methods were similar to those of the first group of projects, including quantitative and qualitative methods such as questionnaire surveys, key informant interviews, focus group discussions and participatory learning activities, knowledge, attitudes, and practice survey, in-depth household interviews and secondary data collection. However, more systematic attention was paid to selection of locations for the study, to ensure any agro ecological or institutional variations were covered, and also to the selection of project and non project villages or respondents for surveys. In several studies these selections were done randomly[98]. These studies tended to combine a range of methods used at different stages and for different aspects of the study. Box 2 describes a qualitative investigation which followed a structured survey.

Box 2 - Rainfed Rabi Cropping in Rice-fallows of Nepal, DFID RNRRS

The qualitative study was conducted in two of the four districts covered by the structured survey, representing high and low adoption of the project outputs respectively. Three villages (out of the six covered in the structured survey) were selected purposively in each district, based on two criteria: the initial level of interest in the new varieties; and the level of social differentiation within the village.
Key informant discussions were used initially to gather generic information on the village and to help to identify and contact the possible different groups to interview subsequently. The focus group discussion was chosen as the preferred tool. Depending on the theme explored, people were selected on the basis of their gender, class or caste or in relation to their professional status. Although the group discussions were fairly open-ended, the survey aimed to cover a number of core topics, such as changes in livelihoods and sustainability of livelihoods.

Source: Conroy, 2009 Institutional innovations in India's crop improvement system: Rainfed Agriculture Impact Assessment Study No. 5 September 2009

Several of the evaluations in this category analysed the economic impacts of technology adoption, examining the yields, income and profits of intervention and control groups and conducted cost benefit analysis. Additional tools used were market surveys, market chain analysis and network mapping.

Some of the projects identified were not one-off evaluations, but were actually projects designed around assessing impact and which monitor the relevant variables which will identify the extent of change. This is especially important where the objective is to understand the impact of the technology or research output on rural livelihoods and the rural poor, for example through tracking the consequences of diversifying income sources, the strength of social safety nets and the impacts on women's work load and role in household decision making.

The programmatic evaluations were similar in methodological approach[99]. Similar types of methods were used, with the addition in one case, of modelling of economic impact, showing the

[97] Projects using a 'before' and 'after' framework and baselines were numbers 33, 35 and 44, table 1.
[98] An example of random selection is number 35, Table 1. Instances of assessment of economic impacts were projects 28, 34 and 35. Projects 16 and 37 combined methods.
[99] The programmatic evaluations in this group were numbers 26 and 49

hypothetical 'without the technology' case and estimating the rate of return, assuming no other changes in the farming system. Modelling did not include externalities, intangibles, and long-term and distributional effects[100].

An ex post analysis of IRRI's contribution to varietal improvement research was interesting in that it traces the genealogies of the varieties released by IRRI and by the NARS, showing that although most of the later releases have come from the NARS, they have as their ancestors an IRRI line. The study also takes into account indirect impacts, for example, increased availability of year-round employment, expansion in agriculturally linked rural non-farm activities, and lower consumer rice prices.

The GTZ funded project 'Assessing the contribution of diversified Musa genetic resources to poverty reduction, environmental sustainability and gender equality in rural communities[101] contains four ex post studies on various aspects of musa germplasm conservation and distribution, testing of cultivars, dissemination of cultivars and their contribution to livelihoods and rural development and the use of impact assessment in improved priority setting and delivery of improved technologies.

Within evaluations in this category there starts to be a discussion of how problems of bias[102] have been addressed – how the evaluation has taken account of the influence of variables outside the planned intervention logic of the project. These could be pre existing systematic differences between farmers in the intervention group and the control group, although these groups are supposed to share similar characteristics and circumstances, or these might arise from the way target groups and comparison groups have been selected (if not randomly and if not appropriately matched). Other influences could arise from the activities of other projects or agencies, or through 'leakage' from the project. The difficulty is that these factors are not always easy to observe.

An example of this is given in the collection of case studies for natural resource management for irrigated rice in Asia[103]. In part three on impact, the case study, '*Three Reductions, Three Gains*' (3R3G) *knowledge-based crop management technology* examines the effectiveness of this crop management strategy for rice. This was aimed at lowering the cost of growing rice in irrigated systems while maintaining yield, improving farmers' health and better protecting the environment. To assess the possible impact of adoption on poverty reduction, survey data was collected on net income from rice, average farm size and household size. Annual per capita incomes were computed for each province and compared to the poverty line.

The study constructed a counterfactual to enable examination of the question of what would have happened to input use, yield, production costs and farm income had IRRI not introduced

[100] Proceedings of Technical Review and Planning Workshop: "Assessing the Impact of Resource Conserving Technologies in the Indo-Gangetic Plain; 5-7 May 2004, Hotel Eastern Residence, Banani, Dhaka, Bangladesh (3); Erenstein 2009 (7).
[101] http://ongoing-research.cgiar.org/factsheets/post-doc-project-assessing-the-contribution-of-diversified-musa-genetic-resources-to-poverty-reduction-environmental-sustainability-and-gender-equality-in-rural-communities/ (51)
[102] Akinola, A.A., Arega D. Alene, R. Adeyemo, D. Sanogo, A.S. Olanrewaju n/d Impacts of Balanced nutrient management systems technologies in the northern Guinea Savanna of Nigeria. Department of Agricultural Economics, Obafemi Awolowo University, Ile-Ife, Nigeria. (31)
http://smartech.gatech.edu/bitstream/handle/1853/35265/1244623341_AA_1.pdf?sequence=1
[103] Palis FG, Singleton GR, Casimero MC, Hardy B, editors. 2010. Research to impact: case studies for natural resource management for irrigated rice in Asia. Los Baños (Philippines):International Rice Research Institute. (27)

3R3G. However, as Box 3 shows, this type of comparison can conceal different types of bias. However, the approaches required to address this (further elaborated in the next section) are currently evolving and are complex, requiring specialist econometric skills.

Box 3. Crop management technology for rice and the pitfalls of random sampling

The study surveyed a random sample of farmers in order to create comparable groups of adopters and non adopters wherein the observed differences in their input use, costs, and incomes were equated to measures of impacts of the intervention. However, just when the data had been analyzed and a report written, the authors learned that the random sampling procedure deployed does not ensure random assignment of "treatment" so there exists the possibility of self-selection into treatment …or adoption decision … relevant to the process determining the outcome (Faltermeier and Abdulai 2009). In other words, the adopters may be systematically different from non adopters in that they may be more knowledgeable and innovative, less capital constrained, less risk averse, and consequently have adopted superior technologies apart from 3R3G. When this happens, there is an upward bias in our impact measures as they also capture the benefits from other technologies the adopters were using concurrently. Solutions to the problem of attribution and creating counterfactuals are becoming commonplace in economic literature as they continue to evolve. But, the econometric procedure is usually difficult to follow and not readily applicable to the specific cases at hand. An impact evaluator therefore needs to keep abreast of the methodological evolution and learn to adapt such to specific impact studies.

Source: extracted from Huelgas and Templeton, 2010[104]

The remaining programmatic evaluations were donor initiated assessments of their contributions. There was an interesting set of reports on the evaluation of the EC contribution to the CGIAR and the practical application of the CGIAR research results by smallholder farmers, including country case studies[105]. There was an evaluation of DFID's renewable natural resources research strategy[106]. There were three reports evaluating Belgian aid (a report with four case study countries; and separate reports on Rwanda and Niger). There was also a joint evaluation of EU AID to Niger from Belgium, France, Denmark and Luxembourg.[107]

4.6.3 Impact assessment – rigorous design

The third group of six project and programme evaluations is characterised by more rigorous design and quantitative methods. The combination of quantitative methods with qualitative and participatory methods, as discussed above, is less emphasised[108]. These are *impact* evaluations and involve a number of design features and statistical techniques to overcome selection bias and more confidently attribute effects (*the average treatment effect*) to programme interventions. Among such techniques is *propensity score matching*, used by five of the six projects[109]. Quasi experimental

[104] Z.M. Huelgas and D.J. Templeton 2010 Adoption of crop management technology and cost-efficiency impacts: the case of the case of Three Reductions, Three Gains in the Mekong River Delta of Vietnam, in Palis et al, 2010 (27)
[105] European Commission 2007, Evaluation of EC contribution to CGIAR 2007 (19) European Commission 2011. Practical Application of CGIAR Research Results by Smallholder Farmers. (20)
[106] Spencer, Dunstan, Stein Bie, Ursula Blackshaw and Anne Thomson 2005, Evaluation of DFID's Renewable Natural Resources Research Strategy1995-2005. DFID EVSUM EV659 (55)
http://www.oecd.org/dataoecd/7/7/35242503.pdf
[107] Project numbers 21, 22, 23, 24 in table 1.
[108] Included in 2 and 11 only
[109] A similar technique 'Nearest neighbour matching' (Abadie and Imbens, 2002) was used by study 11.

approaches are used, comparing 'intervention' and 'control' groups. Randomised experimental designs are rarely used for evaluating developmental impacts of agricultural research, since 'treatments' cannot easily be randomly assigned.

Box 4 - Access, Adoption, and Diffusion; Understanding the Long-term Impacts of Improved Vegetable and Fish Polyculture Technologies in Bangladesh

A survey on the adoption and outcomes of these two technologies conducted in 1996/7 was followed up in 2006/7, allowing for more rigorous evaluation of long term impact of the intervention. The 1996 survey covered households in 'treatment' villages where the technology had been introduced and in comparison villages where the technology was yet to be introduced. Data on production, other income-earning activities, expenditures, food and nutrient intakes, time allocation patterns, and health and nutritional status were collected from adopting households in villages with the technology; from likely adopter households in the villages where the technology was not yet introduced; and a cross-section of all other non-adopting households in the study villages.

The impact evaluation in 2006/7 investigated whether interventions had resulted in improvements in nutritional status, particularly of women and children and the factors underlying the differential impact of the interventions on household and individual-level outcomes. The study used both quantitative and qualitative techniques. The survey covered 957 core households that took part in the original survey and 280 "splits" from the original household. Qualitative methods included key informant interviews, focus groups and life histories to explore participants' perceptions of poverty, livelihoods strategies, the institutional setting, and technology dissemination pathways. Given that the original targeting criteria for the technology were not random, simple comparisons of outcomes between treatment and comparison households would yield biased estimates of program impact. To address this, the study used nearest-neighbour matching to construct a statistical comparison group for the adopting households on the basis of observable household characteristics. This allowed the measurement of the impact of the adopted technologies on households and individuals by comparing actual outcomes for the intervention group with the outcomes for a group of statistically matched households without the interventions. Repeating the survey for both groups 10 years after the original evaluation enabled control for unobservable time-invariant characteristics using difference-in-differences techniques.

The study found that the long-term impacts on household-level consumption expenditures and asset accumulation were, in general, insignificant in the improved vegetables site but positive and significant in the individually operated fishponds sites. Impacts on consumption expenditures and assets were negative and significant in the vegetable technologies sites. In terms of household nutrient availability, impacts were insignificant across the three sites.

Source: Summarised from Kumar and Quisumbing 2010

Another example of designing an impact assessment to overcome selection bias is illustrated in the study of the impact of a farmer field school programme on farmers' knowledge of integrated pest management practices related to potato cultivation in the Peruvian Andes[110]. Household surveys were conducted in thirteen communities where small holder potato farming was the dominant agricultural activity, collecting information on crop cultivation practices, household characteristics, access to services, consumption and farm assets. A knowledge test was conducted based on the curriculum of the farmer field school. The purpose was to measure the impact of FFS on those who participated in the program based on the comparison of FFS participants with non-participants that could serve as good counterfactuals.

[110] Erin Godtland, Elisabeth Sadoulet, Alain de Janvry, Rinku Murgai, and Oscar Ortiz 2003. Testing the Impact of Farmer-Field-Schools on Knowledge: An Empirical Study of Potato Farmers in the Peruvian Andes (12) http://www.3ieimpact.org/admin/impact_evaluations/Impact%20of%20Farmer%20field%20schools%20on%20Knowledge%20and%20Productivity.pdf

CARE had introduced the FFS programme into four villages of the ten in which it was working. The study sample covered these and three control villages which were similar to the FFS villages in observable characteristics such as agro-climatic conditions and infrastructure. To avoid bias from potential diffusion of knowledge within the FFS communities, non-participants in FFS communities were excluded from the analysis. It was assumed that there was no diffusion between FFS and non-FFS villages.

A probability propensity score technique[111] was used to create a control group from farmers in the non-FFS villages, similar to the farmer-field-school participants in observable characteristics. Impact of the programme was estimated by comparing the observed outcome of FFS participants with the outcome of farmers from the comparison group[112]. The study examined the variables at household level which account for the propensity to participate. FFS participation was strongly correlated with the availability of surplus labour in the household, but not wealth nor education. The results indicated no systematic differences between the experiment and control groups, increasing the reliability of the finding that farmers who participate in the program have significantly more knowledge about integrated pest management practices than those in the non-participant control group.

Three additional studies used propensity score matching. The first, 'Impact of Improved Maize Varieties on Poverty in Mexico: A Propensity Score-matching Approach) PSM)'[113] uses a similar methodology to assess the impact of adoption of improved maize germplasm on household welfare in Chiapas and Oaxaca. The authors considered that the use of PSM substantially reduced the risk of bias. The study explored whether the low adoption rates of improved maize germplasm could be explained by low returns. Overall, results suggest that productivity-enhancing agricultural innovations are effective in raising farm incomes, contributing to poverty alleviation and food security. This raises further questions on the reasons for low adoption rates in Mexico and suggests the need for more research on the determinants of technology adoption.

Choice and combination of methods make a difference to results. The study 'Technology adoption, productivity and specialisation of Uruguayan breeders: evidence from an impact evaluation'[114] assesses impact adopting a 'difference-in-difference' with a propensity score matching estimation strategy. They note that there were important differences in results when the 'difference in difference' of beneficiaries and non-beneficiaries is used without Propensity Score Matching (PSM) and when these were both used.

One of the most detailed expositions illustrating the complexity of impact assessment methodology is in the study 'the Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands'[115]. As in the other studies in this group, there is a discussion of the difficulties of constructing a measure of the counterfactual outcome of what production would have been

[111] Rosenbaum, P.R., and Rubin, D.B. (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Biometrika, 70, 41-55.
[112] The methods used for matching were propensity score cut-off point; nearest neighbour matching -including non participants with propensity scores closest to participants' scores; use of the entire non participant sample. A 'balancing test" examined the extent to which the control groups resembled the treatment groups.
[113] The Impact of Improved Maize Varieties on Poverty in Mexico: A Propensity Score-matching Approach, World Development, July 2010, v. 38, 7, pp. 1024-35. (68)
[114] Lopez and Maffioli 2008 (70)
[115] Menale Kassie, John Pender, Mahmud Yesuf, Gunnar Kohlin, Randy Bluffstone, Elias Mulugeta (2007) Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands. IFPRI Discussion Paper 00733, December 2007. International Livestock Research Institute, Environment and Production Technology Division (69) http://www.3ieimpact.org/admin/impact_evaluations/Impact%20of%20Soil%20Conservation%20on%20Crop%20Production.pdf

without conservation on conserved plots, given that randomly assigning plots to treatment and control status in real farming situations was not possible.

Box 5 - Dealing with bias in counterfactuals - Impact of Soil Conservation in the Northern Ethiopian Highlands

Farmers make their own adoption choices, or are systematically selected by development agencies based on their propensity to participate in technology adoption. In addition, farmers (or development agencies) are likely to select plots non-randomly based on their quality attributes, which are often unobservable by the researcher. Therefore, adopters and non adopters may be systematically different, and conserved and non-conserved plots may also be systematically different, and these differences may manifest themselves in differences in farm performance that could be mistakenly attributed to differences in adoption behaviour. Thus, it is difficult to perform ex-post assessment of gains from conservation using observational data, because of possible selection bias due to observed and unobserved plot and household characteristics. Failure to account for this potential selection bias could lead to inconsistent estimates of the impact of technology adoption.

Source: extracted from Kassie et al 2007:6.

The study collected data from more than 900 households and plots of conserved and non conserved land in Amhara and Tigray, Ethiopia in 1999 and 2000. Several methods were used to check for selection bias or endogeneity problems that could arise due to correlation of observed factors with unobserved variables. Propensity score matching was used to select matched samples of conserved and non-conserved plots[116]. The nearest neighbour matching method was also used to generate the samples based on observable variables. Regression and stochastic dominance analyses were based on these matched samples.

At a programme level, the Sub-Saharan Africa Challenge Programme (SSACP) has designed its impact assessment to provide 'proof of concept' of Integrated Agricultural Research for Development, IAR4D. This is different from the mainly single technology or commodity based impact assessments available in the literature since the IAR4D approach emphasises integrated approaches across value chains, establishing broader partnerships and 'innovation platforms' (IPs), strengthening participation, building linkages with policy processes and stimulating institutional change.

Impact assessment for SSACP is intended to test and validate the conditions under which IAR4D works and whether it delivers greater impact and is more cost effective than traditional approaches. Baseline studies were designed and carried out in 2008 and a database of process and impact indicator variables for the IPs and their associated research communities and households was developed. The main comparisons will be between IP intervention villages and counterfactual comparison villages and households. There was further analysis[117] of the baseline data from the SSACP West Africa pilot learning site, to show whether the households in the baseline study fall into three distinct groups for implementation of the IARD as planned in the IA design. The three groups are IARD farmers, conventional farmers or clean sites farmers'. This linear discriminant analysis indicated an overall rate of 99% of farmers correctly classified into their respective sites.

[116] Kassie et al 2007 (69)
[117] Mokwunye and Ellis-Jones 2010 (2)

The different villages chosen for the program evaluation were also correctly identified within the defined categories.

However, these methods are only partially adequate for the complex relationships in IAR4D since the nexus of partnerships and institutional interrelationships within specific innovation platforms cannot be replicated as a treatment[118]. The programme's intention of learning across cases or the different pilot sites recognised this reality. Furthermore, as IAR4D principles become more mainstreamed in associated research networks (e.g. ASARECA) there are fewer examples of wholly 'conventional' research approaches and more frequent cases of partial and variant integration of IAR4D principles in particular institutional contexts. Comparisons are relatively easier at the level of research sites but less straightforward at the level of institutional change and participation in innovation platforms. Here, complementary methods are needed and the systematic design, presentation and justification of these represent a current practical challenge[119].

The combination of quantitative *and* qualitative methods in impact evaluations are more likely than either type of method alone to be able to shed light on *why* innovation and its impacts are distributed in certain ways, as well as measuring their extent.

Generally few of the studies appear to draw on recent developments in debates on rigorous impact assessment or alternatively, on participatory or narrative methods exploring a theory of change. We have noted that theory of change methods have comparatively recently gained a higher profile, particularly in evaluations of complex relationships. They are relatively easy to introduce at any stage of a research process. In contrast, rigorous evaluation methods are most effectively designed at the outset of an ARD process; they require specialised statistical input and the expense may be difficult to justify within a short planning and funding time frame. Thus they are more likely to be part of a larger programme (e.g. SSA Challenge Programme) than smaller research projects.

Rigorous evaluations can provide important information on the extent of impacts, but these are generally measured at household level. The EIARD task force recommendation to explore the complex social, economic, political and institutional dimensions is not reflected in these approaches. Better collection and utilisation of monitoring data would be helpful, since an understanding of the processes of project delivery, changing relationships and stakeholders' perspectives is important in interpreting results from impact studies.

[118] Martin, Adrienne, 2009. So what difference does it make? Assessing the outcomes and impacts of farmer participatory research. In Ian Scoones and John Thompson eds., Farmer First Revisited; Innovation for Agricultural Research and Development, Practical Action, pp276-281
[119] Ibid.

5 Communication and utilisation of evaluation findings

This section examines available information on how evaluations are used. It looks at how evaluation results are shared among key stakeholders and through what channels. Secondly, it examines how clients and users have utilised the evaluation findings within their own organisations and how findings have been used externally, beyond the immediate client. The difficulty in exploring these questions has been the lack of sources of information on dissemination and communication activities following impact evaluations. This is probably because evaluation studies conducted by consultants are submitted to a donor, who is then responsible for the follow-up. Many of the project evaluations are produced as grey literature reports or journal articles, but there is limited gathering of these dissemination outputs together to be made available on the web linked to the original evaluation reports or analysis of how donors have responded.

5.1 Dissemination of evaluation findings

Few of the studies indicate in their methodology sections how they will seek to disseminate the findings to different audiences. There is information on mechanisms and recommendations to promote the uptake of the technologies but little on the dissemination and utilisation of the evaluation findings.

The users of the findings of impact evaluations and the channels through which they will be reached are not well defined. No information on plans or actual activities for disseminating or communicating the evaluation results could be found for 31 of the 44 project and programme evaluations reviewed.

Few of the databases adequately address the dimension of communications in general and do not have sections dealing with evaluation or dissemination of evaluation findings. More of this information would enable greater evaluation of the success of impact evaluations themselves, as well as providing insights into specific successes (or failures) associated with donor agricultural research for development.

An important ethical question arises as to how far downward accountability (to different stakeholders, national government representatives, private sector people local communities) is served by the types of impact evaluation conducted by donors as much as upward accountability (e.g. to decision-makers). How far do researchers communicate their findings adequately back to those who have participated in the research? If they do not, is this because of a lack of adequate resources and priority being placed upon this by the client, or because they themselves do not recognise the importance of this aspect of the work.

No evidence could be found of creative methods (e.g. community or interactive radio, participatory video) for promoting dialogue between end beneficiaries and those commissioning evaluations. The use of non-written materials is somewhat critical if findings are to be shared to non-literate audiences. Given the lack of information on dissemination and communication it was also difficult to judge the extent to which findings are routinely translated into other languages to more widely share the insights and lessons of work undertaken.

In several cases it is difficult to distinguish the dissemination of project and programme results from the dissemination of evaluation findings. The main dissemination outputs are published articles on the technologies. Dissemination events held at the end of projects and programmes often cover issues in technology development and assessment as well as evaluations of performance and impact. The engagement of different stakeholder groups such as NGOs, policy makers and the private sector in project processes and project evaluations creates a link for communication of the evaluation findings and further dissemination of the technologies.

5.2 Donor feedback on evaluations

There are very few examples of donor responses alongside independent impact evaluation reports and articles. Donor responses are an important part of the picture, providing an opportunity for the donor to comment on the findings and to indicate how they might change their practices and policies as a result and what they intend to not change. This phase in impact evaluation is becoming more commonly practiced by clients and is an important part of the evaluation process.

DFID's response to their externally conducted impact evaluation of the Renewable Natural Resources Programme by Spencer *et al* [120] stated;

'*The findings and recommendations of the evaluation are timely as DFID prepares for a new Strategy for Research on Sustainable Agriculture (SRSA). A document on the SRSA draws on the lessons learned from the evaluation and is the current (June/July 2005) subject of widespread consultations. These lessons, in particular the need for stronger management in some areas, coherent monitoring and evaluation systems and a smooth transition from the RNRRS to the SRSA will be built into the design of new programmes which DFID expects to commission later in 2005/06*',

This kind of clear indication of how lessons will be taken up is quite rare in the studies reviewed and in associated documents publicly available on the web[121].

5.3 Utilisation of evaluation findings

There are current approaches in evaluation which emphasise utilisation[122]. From this perspective, evaluations should be designed working with the intended users who will apply evaluation findings and implement recommendations.

The extract from the DFID evaluation in the previous section indicates how evaluation findings can help to inform the design of future programmes. Other possible uses of evaluations are outlined in Table 7, distinguishing between uses of the results and uses of the evaluation processes.

While many of the evaluations reviewed potentially contribute in the areas defined in Table 7, there was little information available as to how those contributions had been taken up and implemented to serve the accountability function, for learning and improvement or to influence policy.

[120] Spencer et al, 2005 p3 (55)
[121] Some organisations that commission impact assessments, such as the Fairtrade Foundation or Fairtrade International, as a matter of practice produce a response for publication at the same time as the impact evaluation is made available. This is to help protect their reputation but also to indicate how they intend to follow up to a wider external audience.
[122] Patton, 2008

Table 7 - Utilisation of evaluation findings

Uses of findings and results	Use of evaluation processes
To see what has been achieved and to justify funding by clients	To improve communication, information and the relationship between clients and research organizations
To identify strengths and weaknesses, learn from errors, feedback information into planning and improve the research approach	To create an environment of critical self-reflection and a culture of learning
To ascertain cost effectiveness	To empower clients (encourage clients through 'downward accountability')
To generate knowledge and share lessons and new concepts inside the system	To generate knowledge together and share lessons and new concepts with those outside the system
To influence policies and sectoral priorities	

Source: Adapted from Hoffmann *et al.* 2009, (in Martin *et al.*, 2011)

Only 5 EIARD focal points (DFID, EC, Germany, Italy and Switzerland), responded to the question on the ways in which Impact Assessments have been used internally and externally.

DFID RNRRS impact assessments were used for internal learning and for planning future resource allocations. The Evaluation of EC contribution to CGIAR (2007) and Practical Application of CGIAR research results by Smallholder Farmers (2011) were used internally to justify continued EC support for CGIAR and areas of strategic focus. They were externally shared as part of the EC input to broader European coordination of support to CGIAR.

Germany's evaluations of GTZ were used to inform their own programmes while externally the use of the results was up to the CGIAR programme.

Italy's evaluations of their funded programmes are used internally for programming of following year's contributions to ARD programmes and/or research centres. Their ARD Impact Assessment is currently being used to develop Rural development, Food Security and Agriculture guidelines.

Switzerland reported that the various SDC evaluations were used in different ways. The findings of the 5 Year Ex-Post Impact Study of the POSTCOSECHA Programme Central America were used externally to promote the approach within the thematic network. The evaluation of the Irrigated Rice Research Consortium in South and South East Asia was used for internal institutional learning for IRRC and IRRI. The main external influence was through the programme's publications, workshop contributions and seminars.

SDC's broader sectoral evaluations had more strategic uses, both internally and externally. The evaluation of SDC contribution towards biodiversity: impact in the Andean region was used

internally by senior management response and the resource person for Biodiversity Dissemination in the Swiss administration, while externally the findings were shared through DAC EvalNet. The Evaluation of SDC's Research Related Activities was used by management to shape a new institutional set up as well as dissemination through DAC EvalNet.

The study commissioned by the Swiss Development Cooperation to assess the effectiveness of its research provided recommendations for improvement[123]. These were to develop a new research policy and the organisational structures required to implement it; to define and adopt "essential standards" for results based research management and adapt existing information systems to facilitate strategic oversight, research project management, knowledge management and communication of research results and to develop mechanisms to maximize the use of the results of research, including within SDC's own operations. The focal point reported that it was the basis of accountability to parliament and was used for setting future developmental priorities.

However, we do not have information as to how the findings and recommendations actually shaped SDC policy and practice. This would require in-depth country case studies and face to face meetings as such policy informing and influencing processes do not appear to be documented. It was difficult to find policy briefs and summaries associated with any of the impact studies.

With the exception of knowledge sharing, the 'process' uses of evaluations in Table 7 were not mentioned.

An important observation relating to the content of the evaluations is that negative outcomes and impacts are rarely reported, despite the OECD-DAC definition of impact as including both positive and negative effects. This limits the extent to which evaluations can provide credible evidence on which to base improved design and delivery of future research.

Relating these observations to the discussion above on methodology, it appears that the rigorous impact evaluations focusing on particular aspects of ARD and technology development are useful for assessing the cost effectiveness of investments, and for sharing knowledge on technology performance and upscaling. The broader type of programmatic evaluations which more comprehensively address a range of issues including influences on research capacity and institutions, appear to be more influential in terms of decision making on policy and future investment. This would suggest that focussed rigorous impact assessments can complement, but cannot replace evaluation approaches which systematically explore the *plausible* linkages between development research and its outcomes and impacts across different dimensions. Stakeholders associated with these two perspectives, in particular the research organisations and donors, could both benefit from supporting complementary approaches.

[123] The Policy Practice Limited, UK 2009, Long evaluation abstract, in SDC 2010, Evaluation of SDC's Research Related Activities. (29)

6 Challenges and Opportunities

6.1 Methodological challenges and strengthening good practice

The discussion of evaluation designs and methods in section 4.6 above indicated the tendency for evaluations which are broad in scope and combine some quantitative methods with qualitative and participatory approaches to be less rigorous in measuring and attributing impact. Conversely, those which were more rigorous (a far smaller numbers) were generally not able to address holistically 'the complex social, economic and political dimensions of pro poor innovation'. Other studies have confirmed this finding. A review of evaluations of farmer field schools[124] found that the studies were designed to be either statistically rigorous (but with a restricted scope) or comprehensive (but with limited coverage), but never both, which had negative consequences for their overall conclusions on performance.

There is a need for more 'hybrid' methodologies[125] using elements of participatory and conventional approaches for evaluating research and development programmes. This requires incorporation of this awareness into the drafting of terms of reference for evaluations and team skill specifications.

It has been observed by an Evaluation Gap working group, in their document '*When will we ever learn*', that rigorous evaluations of social programmes are relatively rare, partly because they fall outside normal budget and planning cycles, there are too few incentives and too many obstacles[126]. A similar situation exists for evaluation of ARD.

Ways to address this include more collective commitment (donors and national governments) to better coordination and joint funding of impact evaluations across countries and institutions around common thematic areas and for governments and agencies to reinforce efforts to generate and apply knowledge from impact evaluations. The importance of strengthening overall monitoring and evaluation systems is also stressed.

6.2 Prospects for greater coordination and harmonisation of IA efforts

It was difficult to get a detailed insight into this question given the study limitations and reliance on documentation rather than interaction with the main EIARD decision makers. However, some observations are presented.

The prospects for greater coordination and harmonisation of impact assessment efforts are influenced by several factors. These include the purposes and expectations of evaluations and the methodologies utilised and the different funding scenarios of EIARD members. Also of significance are the different institutional responsibilities for evaluation and the legal position (for example, the EC Financial and Administrative Framework with International Organizations).

In terms of funding to the CGIAR, for countries who provide all, or most of their funding through the CGIAR in unrestricted form, the expectation of the CGIAR is that they will rely on

[124] van der Berg (2004)
[125] Njuki *et al.* (2008)
[126] Center for Global Development (2006)

the Independent Evaluation Arrangement (IEA) to commission and conduct evaluations, following consultation on the terms of reference and major issues to include in the evaluation. The policy[127] states that evaluations should be responsive to issues of major stakeholders, including donors and partner country governments. This consultation process would ensure that the views and interests of European donors would be represented.

Where funding to the CGIAR is restricted, or where the donor requires their own evaluation, then joint evaluations 'will be considered'. These would have to be worked out with the IEA.

For the funding for ARD outside the CGIAR, there may be potential for joint evaluations with a more thematic focus. The responses and documents from EIARD members did not indicate any good examples of where this currently exists, however, there has been interest expressed in the past[128]. Examples of possible themes are capacity building for ARD, innovation systems approaches to research.

The institutional modalities for evaluation and impact assessment are topics of current discussion. What is clear is that impact assessment in the sense of measuring attribution, utilising rigorous and statistically sophisticated methods is a specialised function and may be best contracted independently.

Evaluation, according to the evaluation standards is an independent function, '*Evaluators are independent from the development intervention, including its policy, operations and management functions, as well as intended beneficiaries*'[129]. On the other hand evaluation can be seen as an internal function linked with monitoring and learning. There are trends in institutional practices aimed at separating accountability functions from wider evaluation purposes. Examples here are UK, EC and Sida, the latter having an Aid Watch Dog for accountability purposes and an internal evaluation department focusing on utilization-based evaluation with a strong learning orientation.

[127] CGIAR Fund, 2012, CGIAR policy for Independent External Evaluation, p 7
[128] Ferreti et al, 2008.
[129] OECD. 2010.

7 Conclusions

This paper has outlined some of the methodological challenges in evaluation and impact assessment and examined the available materials from EIARD members to assess their policies and practices.

In comparison to the recommendations of the EIARD task force paper 2001, the findings relating to evaluations commissioned by EIARD members, suggest that specification of the subject matter of the research and its contextualisation, the objectives and scope of the evaluation and discussion of other influencing factors, are well covered in current practice. However, other important recommendations are followed in a minority of cases. These include, making explicit the model or concept of innovation, explaining the logic model underlying the programme or project and a statement or hypothesis of the impact that is expected. Although not explicit in most documents, the recommendation that a plan for impact assessment and evaluation should be prepared before the project commences and be an integral part of project implementation does not appear to have been implemented. Finally, there is scope for much wider inclusion of critical review and comment from different stakeholders, partners and beneficiaries.

Most of studies and reports reviewed are project and programme evaluations. There were few impact assessments involving measurement of actual changes and attribution. The documents reviewed are those which were made accessible to the study team or are available on- line, so may not represent the evaluation output as a whole. In addition, there was underrepresentation of more recent studies. Nevertheless, it appears that many of the evaluations did not acknowledge either recent developments in debates on rigorous impact assessment, nor on the other hand, include participatory or narrative methods. Many studies were not clear whether or not baselines were present.

Evaluations are often limited by the restricted depth and coverage of monitoring data which could help to substantiate the pathway towards impact and enable understanding of processes of project delivery as they emerge.

There is generally a lack of disaggregation of data in evaluations to indicate what kind of impacts and for whom. In particular there was limited identification of gender and poverty related impacts.

Considering the high level of direct support to the CGIAR, there is relatively little engagement in evaluation processes or utilisation of the CGIAR impact assessments. Only the EC appears to have examined the impact of CGIAR outputs. It is not clear whether other countries funding CGIAR directly had used these reports to inform their own decision making.

7.1 Recommendations

There are a number of recommendations which could enhance the efficiency and effectiveness of evaluation and impact assessment among EIARD members.

- Good evaluation and impact assessment begin with project design. It is important to develop impact oriented thinking, and to encourage the inclusion of evaluation plans and

IA design in the project design and implementation plans. Focused baseline information collection can greatly enhance the capacity to assess outcomes and impacts.

- More emphasis on effective monitoring could encourage understanding of processes and achievements of ARD projects as they emerge. EIARD should engage with and support the current CGIAR efforts to develop more robust monitoring and evaluation systems in addition to the formal external ex post impact assessments conducted.

- There is a need to build understanding amongst those commissioning evaluations of ARD of the different kinds of evaluation and impact assessment and to guide choices in design and methods to be appropriate for specific objectives and circumstances.

- In commissioning evaluations, the expectations and type of evaluation required should be made clear. Terms of reference need to clearly specify the purpose of the evaluation and what is actually required. This is the basis for determining choice of methods.

- The planning of an evaluation should include a clear timeline, a step for critical review and comment from different stakeholders and a plan of action for communication of the evaluation findings.

- There is a need for development and agreement on procedures to encourage the sharing and dissemination of evaluation findings among EIARD members and their wider stakeholders. To help harmonize consistency and quality of reporting for ARD evaluations a best practice guide on quality standards specifically for ARD could be developed for EIARD members.

- Improvements to the ARD databases could increase the accessibility of evaluation and impact assessment reports. Evaluation reports and impact assessments should be categorised more clearly. Data bases could incorporate 'evaluation' as a search theme; include fields on dissemination of the evaluation findings; provide URLs to share reports on the website, and enhance the narrative descriptions of what has been done. With this additional information the research databases could extend understanding of the impact evaluation process, to include more information on the process of framing research questions (by whom, how, when, why) and of commissioning evaluators (how, why, who etc), as well as the post-study phase of absorbing the findings and the political process of responding and internally deciding what to do with the results.

- Apart from improvements to existing data bases, EIARD members should explore their joint willingness to establish a web site or web page for open sharing of evaluation reports.

- Greater interest and commitment to develop joint studies should be encouraged to enhance methodological rigour and shared learning.

- For those conducting evaluations and impact assessments, there is a need to develop guidance for impact evaluation planning which helps in the selection of evaluation approaches appropriate for complex situations. The specific tools and techniques used should be consistent with the principles underpinning the evaluation and its objectives and tailored to facilitate exploration of the evaluation questions within the time and resources available.

- Multiple methods are preferable, exploring the meaning and the measurement of project impacts. There is scope to innovate and support participatory, qualitative and mixed-

methods, combining and sequencing different approaches and tools in evaluation In this way, evaluations can explore the impacts of what has been done as well as the strategic and institutional positioning of ARD interventions.

- The development and use of flexible and non-linear programme theories of change should be incorporated as a standard tool within evaluation and specifically required in terms of reference. These take into consideration other actors and processes often neglected by logframes and linear impact pathways. They can be developed with the participation of stakeholders. They can help to explain the logic underlying the programme or project and to define the impact that is expected.
- The impact pathways should seek to disaggregate impacts for different stakeholder groups and in particular identify gender and poverty related impacts.
- Rigorous and quasi experimental approaches can be useful for assessing impact of specific sub-components of projects, particularly for technology components. They are less suitable for the complex, interactive, multi-stakeholder approaches of ARD.

8 References

Abadie and Imbens, 2002 Simple and Bias-Corrected Matching Estimators for Average Treatment Effects, Technical Working Paper T0283, National Bureau of Economic Research http://www.nber.org/papers/t0283.pdf

Asian Development Bank 2005. An impact evaluation of the development of genetically improved farmed tilapia and their dissemination in selected countries. Operations Department (36). http://www.extension.org/mediawiki/files/f/f5/An_Impact_Evaluation_of_the_development_of_Genetically_Imprv.pdf

Behailu, M, M Abdulkadir, A Mezgubu and M Yasin 2005. Community based irrigation management in the Tekeze Basin: Performance Evaluation: A Case study on three small scale irrigation schemes (micro dams). (1)
http://www.iwmi.cgiar.org/assessment/files/word/ProjectDocuments/ILRI/Atinkut_Community%20Based%20Irrigation%20_ProjectReport.pdf

CABI Bioscience, 2004. Proceedings of Technical Review and Planning Workshop: "Assessing the Impact of Resource Conserving Technologies in the Indo-Gangetic Plain; 5-7 May 2004, Hotel Eastern Residence, Banani, Dhaka, Bangladesh. http://www.dfid.gov.uk/r4d/SearchResearchDatabase.asp?OutputID=173983.

Center for Global Development, 2006. When Will We Ever Learn? Improving Lives through Impact Evaluation. Report of the Evaluation Gap Working Group. May 2006. http://www.cgdev.org/files/7973_file_WillWeEverLearn.pdf

CCIAR, 2009. Voices for Change. The new CGIAR.
http://www.cgiar.org/pdf/Business%20Meeting%202009/voices_for_change_final_jan2010.pdf

CGIAR 2011 Joint declaration, annex Conduct.
http://www.cgiar.org/pdf/Business%20Meeting%202009/jointdeclar_final_jan2010.pdf

CGIAR 2011 A strategy and results framework for the CGIAR, February 20, 2011. http://www.cgiarfund.org/cgiarfund/sites/cgiarfund.org/files/Documents/PDF/srf_feb20_2011.pdf

CGIAR 2011 CGIAR Standing Panel on Impact Assessment: Strategy and Operational Plan 2011-13 http://impact.cgiar.org/sites/default/files/images/SPIAstrategy2011-13.pdf

CGIAR 2012. CGIAR policy for Independent External Evaluation. CGIAR Fund, January 2012.

Cuellar, Melinda, Hans Hedlund, Jeremy Mbai, Jane Mwangi 2006. The National Agriculture and Livestock Extension Programme (NALEP) Phase I Impact Assessment. Sida Evaluation 06/31. (13) http://www.sida.se/Documents/Import/pdf/0631-The-National-Agriculture-and-Livestock-Extension-Programme-NALEP-Phase-I-Impact-Assessment.pdf

Czech Conroy, 2009 Institutional innovations in India's crop improvement system: Rainfed Agriculture Impact Assessment Study No. 5 September 2009 Research into Use. http://www.researchintouse.com/resources/riu09impact5india-crop.pdf

DANIDA 2006 Evaluation Guidelines. Copenhagen: DANIDA. http://www.netpublikationer.dk/um/7571/ (accessed 3 January 2012).

Davies and Dart 2005. The most significant change (MSC) technique: A guide to its use. http://www.mande.co.uk/docs/MSCGuide.pdf (accessed 3 January 2012).

de Janvry, Alain, Andrew Dustan, and Elisabeth Sadoulet. 2011. 'Recent Advances in Impact Analysis Methods for *Ex-post* Impact Assessments of Agricultural Technology: Options for the CGIAR' Report prepared for the CGIAR. April 2011 http://areweb.berkeley.edu/~sadoulet/papers/deJanvryetal2011.pdf

DFID, 2009. Building the evidence to reduce poverty The UK's policy on evaluation for international development. Department for International Development (DFID) June 2009 http://www.dfid.gov.uk/Documents/publications1/evaluation/evaluation-policy.pdf http://www.oecd.org/dataoecd/39/54/43183649.pdf

Duflo, Esther, Glennerster, Rachel, and Kremer, Michael, "Using Randomization in Development Economics Research: A Toolkit" (2006). MIT Department of Economics Working Paper No. 06-36.

Duignan, 2009 Selecting impact/outcome evaluation designs: a decision-making table and checklist approach. Outcomes Theory Knowledge Base Article No.256. http://knol.google.com/k/paul-duignan-phd/selecting-impactoutcome-evaluation/2m7zd68aaz774/115 (accessed 3 January 2012).

Earl, S., F. Carden, and T. Smutylo. 2001. Outcome Mapping: Building Learning and Reflection into Development Programs. Ottawa: IDRC. http://www.idrc.ca/en/ev-9330-201-1-DO_TOPIC.html (accessed 3 Jan 2012).

Erenstein, O. 2009, Zero Tillage in the Rice-Wheat Systems of the Indo-Gangetic Plains A Review of Impacts and Sustainability Implications IFPRI Discussion Paper 00916 November 2009 http://www.ifpri.org/sites/default/files/publications/ifpridp00916.pdf

Eyhorn, Frank, Paul Mäder, Mahesh Ramakrishnan 2005, The Impact of Organic Cotton Farming on the Livelihoods of Smallholders Evidence from the Maikaal bioRe project in central India (48) http://www.fibl.org/fileadmin/documents/en/development-cooperation/production-systems/executive_summary.pdf

European Commission 2008 Guidelines on Agricultural Research for Development http://ec.europa.eu/development/icenter/repository/B2_EC_Guidelines_ARD_june08.pdf

European Commission. 2004. *Evaluating EU Activities – A Practical Guide for the Commission Services.* Brussels: European Commission. OECD. http://wiki.interact-eu.net/display/pc/0.+Practical+Handbook+for+Ongoing+Evaluation+Contents

European Commission 2011. Practical Application of CGIAR Research Results by Smallholder Farmers.

EIARD Strategy 2009–2013. November 2008 http://www.eiard.org/media/uploads/documents/strategies/eiard_strategy_2009-2013_final.pdf

EIARD A strategy for the European Initiative for Agricultural Research for Development (EIARD 1) 2005-2010

EIARD 2007 Epilogue. *Impact Assessment and Evaluation in Agricultural Research for Development.* Task Force on Impact Assessment and Evaluation, European Initiative for Agricultural Research for Development (EIARD), Henning Baur (Germany), Guy Poulter (UK), Marina Puccioni (Italy), Pilar Castro (Spain), Hans-Jörg Lutzeyer (EIARD) and Stephan Krall (Germany).

http://www.eiard.org/media/uploads/File/documents/impact_assessment_paper-1(1).pdf

Ferreti, E.L, Martin A.M. Tollens, E. 2008. Towards Common Donors' Approaches for the Monitoring and Evaluation of the CGIAR – A European Initiative. Task B Monitoring of the European Commission support to strategic Agricultural Research through the CGIAR centres.

Fischler, Martin 2011. 5 Year Ex-Post Impact Study on the POSTCOSECHA Programme Central America. Intercooperation.

http://www.postcosecha.net/en/Home/Documentation/Publications/Publications_in_English

Gabre-Madhin, Eleni, Marcel Fafchamps, Richard Kachule, Bio Goura Soule, Zahia Kahn 2001. Impact of agricultural market reforms on smallholder farmers in Benin and Malawi. Final

Report Volume 2. Submitted to the Deutsche Gesellschaft Für Technische Zussammenarbeit (GTZ) (42) http://www.ifpri.org/sites/default/files/publications/v2.pdf

GTZ, 2008 Results-based Monitoring Guidelines for Technical Cooperation GTZ Corporate Development Unit 2008
http://www.csr-weltweit.de/uploads/tx_jpdownloads/wirkungsorientiertes-monitoring-leitfaden-en_01.pdf

Godtland, Erin, Elisabeth Sadoulet, Alain de Janvry, Rinku Murgai, and Oscar Ortiz 2003. Testing the Impact of Farmer-Field-Schools on Knowledge: An Empirical Study of Potato Farmers in the Peruvian Andes.
http://www.3ieimpact.org/admin/impact_evaluations/Impact%20of%20Farmer%20field%20schools%20on%20Knowledge%20and%20Productivity.pdf

Golam Faruque, 2007 An exploration of impacts of aquaculture production and marketing on rural livelihoods in three regions in Bangladesh. Thesis submitted for the degree of Doctor of Philosophy. University of Stirling. (44)
https://dspace.stir.ac.uk/bitstream/1893/253/1/PhDThesis_GolamFaruque.pdf

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

Huelgas Z.M. and D.J. Templeton 2010 Adoption of crop management technology and cost-efficiency impacts: the case of the case of Three Reductions, Three Gains in the Mekong River Delta of Vietnam; in Palis FG, Singleton GR, Casimero MC, Hardy B, editors, 2010. Research to impact: case studies for natural resource management for irrigated rice in Asia. Los Baños (Philippines): International Rice Research Institute.
http://books.irri.org/9789712202599_content.pdf

James Morton & Co. 2010, Analysis of donor support to CAADP Pillar 4. Draft final report to the European Initiative for Agricultural Research for Development (EIARD)

Kassie, Menale, John Pender, Mahmud Yesuf, Gunnar Kohlin, Randy Bluffstone, Elias Mulugeta (2007) Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands. IFPRI Discussion Paper 00733, December 2007. International Livestock Research Institute, Environment and Production Technology Division.
http://www.3ieimpact.org/admin/impact_evaluations/Impact%20of%20Soil%20Conservation%20on%20Crop%20Production.pdf (69)

Kelley, T., Ryan, J. and H. Gregersen (2008) Enhancing ex-post impact assessment of agricultural research: the CGIAR experience. *Research Evaluation*, 17 (3), 201-212
http://impact.cgiar.org/sites/default/files/images/Kelleyetal2008.pdf

Kumar, Neha and Agnes R. Quisumbing 2010, Access, Adoption, and Diffusion; Understanding the Long-term Impacts of Improved Vegetable and Fish Technologies in Bangladesh, Poverty, health and nutrition division IFPRI Discussion Paper 00995 June 2010
http://www.3ieimpact.org/admin/impact_evaluations/Access,%20Adoption,%20and%20Diffusion.pdf (11)

Kürschner, Ekkehard, Christian Henschel, Tina Hildebrandt, Ema Jülich, Martin Leineweber, Caroline Paul 2010 Dissemination, Adoption and Short Term Impacts of Alternate Wetting and Drying (AWD) in Bangladesh. Dhaka/Berlin (28). http://edoc.hu berlin.de/series/sle/241/PDF/241.pdf

Leeuw, Frans and Jos Vaessen, 2009. Impact Evaluations and Development. Nonie Guidance on Impact Evaluation. NONIE - The Network of Networks on Impact Evaluation.
http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf

Lopez, Fernando and Alessandro Maffioli 2008. Technology adoption, productivity and specialisation for Uruguayan breeders: evidence from an impact evaluation. Inter-American Development Bank Office of Evaluation and Oversight, July 2008, Working Paper. (70) http://www.3ieimpact.org/page.php?pg=details_impact&id=149

Maredia, M. and D. Raitzer 2006. CGIAR and NARS partner research in sub-Saharan Africa: evidence of impact to date, Science Council Secretariat, Rome. http://www.fao.org/docs/eims/upload/258389/CGIAR_WP_AI-proof3.pdf

Martin, A.M, 2009. So what difference does it make? Assessing the outcomes and impacts of farmer participatory research. In Ian Scoones and John Thompson eds., *Farmer First Revisited; Innovation for Agricultural Research and Development*, Practical Action, pp276-281

Martin, A.M. Sabine Gündel, Essie Apenteng and Barry Pound 2011. Review of literature on evaluation methods relevant to extension. http://www.g-fras.org/en/knowledge/gfras-publications/file/63-review-of-literature-on-evaluation-methods-relevant-to-extension

Mokwunye, Uzo and Jim Ellis-Jones, 2010. SSA Challenge Programme, Internal Review Report. FARA, August, 2010.

http://www.dfid.gov.uk/r4d/projectsAndProgrammesResults2.asp?search=related%20Documents&Outputs=Yes&projectID=60686

Morton, J.F. 2010, The innovation trajectory of sleeping sickness control in Uganda: research knowledge in its context. Discussion Paper 09. Research into Use. http://www.researchintouse.com/resources/riu10discuss08ssickcntrl-ug.pdf

NORAD. 2006. Evaluation Policy 2006–2010. Part 1 Strategic priorities, Part 2 Evaluation Programme 2006–2008, Part 3 Guidelines for Evaluation of Norwegian Development Cooperation. Oslo: NORAD. http://www.norad.no/en/evaluation/handbook-and-reference-documents/_attachment/106229?=true&_ts=11eb62db87a (accessed 3 January2012).

Njuki, J., M. Mapila, S. Kaaria, and T. Magombo. 2008. Using community indicators for evaluating research and development programmes: Experiences from Malawi. *Development in Practice* 18(4):633–642.

OECD-DAC, Glossary of Key Terms in Evaluation and Results, Paris, 2002. DAC, http://www.oecd.org/dataoecd/29/21/2754804.pdf

OECD 2010, Quality Standards for Development Evaluation. http://www.oecd.org/dataoecd/55/0/44798177.pdf

Ooijen, Rudy, and David Coombs 2007. Evaluation of EC Contribution to the Consultative Group on International Agricultural Research (CGIAR), Final Report (Volume I) ECORYS Nederland BV

Palis FG, Singleton GR, Casimero MC, Hardy B, editors. 2010. Research to impact: case studies for natural resource management for irrigated rice in Asia. Los Baños (Philippines):International Rice Research Institute (27) http://books.irri.org/9789712202599_content.pdf

Patton, M.Q, 2008 *Utilization-Focused Evaluation.* Thousand Oaks: Sage Publications.

Patton, M.Q. 2010. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. New York, NY: Guilford Press.

Pemsl, Diemuth E. and Terence Too, June 2008. Determination of high-potential aquaculture development areas and impact in Africa and Asia (38)

Pound, B, Michiel van Dijk, Yuca Waarts and Essie Apenteng 2011, Making ARD more pro-poor; Improving accessibility and relevance of results to the poorest. Agrinatura.

Preskill, H., & Coghlan, A. (2003). Evaluation and appreciative inquiry. *New directions for evaluation*, Vol. 100. San Francisco: Jossey-Bass.

Raitzer DA. 2003. Benefit-Cost Meta-Analysis of Investment in the International Agricultural Research Centres of the CGIAR. CGIAR Science Council Secretariat, Washington, and FAO, Rome. 45 p. http://impact.cgiar.org/pdf/126.pdf

Raitzer, David A and Timothy G Kelley 2008, Assessing the contribution of impact assessment to donor decisions for international agricultural research, *Research Evaluation,* 17(3), September 2008, pages 187–199.
http://impact.cgiar.org/sites/default/files/images/RaitzerKelley2008.pdf

Renkow, M and D. Byerlee (2010) The impacts of CGIAR research: A review of recent evidence. *Food Policy* Volume 35, Issue 5, October 2010, Pages 391-402

Research into Use 2006 Learning from the Renewable Natural Resources Research Strategy Pathways for change: monitoring and evaluation
http://www.dfid.gov.uk/r4d/pdf/ThematicSummaries/Brief3_Pathways_for_change_monitoring%20and%20evaluation.pdf

Roche, C, 1999, Impact Assessment for Development Agencies: Learning to value change. Oxford Oxfam/Novib.

Rogers, E.M. 2003. Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press

Rosenbaum, P.R., and Rubin, D.B. 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika,* 70, 41-55.

Rothschild, G. and E. Tollens, 2011. New Terms of Reference for Joint Review Missions applicable in 2011 and 2012 in the EIARD Framework for the Monitoring & Evaluation of the CGIAR. IFAD – AGRINATURA-EEIG.

Sida. 2004. *Looking Back, Moving Forward. Sida Evaluation Manual*. Stockholm: Swedish International Development Cooperation Agency (Sida).
http://www.alnap.org/pool/files/evaluation_manual_sida.pdf

Sidersky P and Guijt I. 2000. Experimenting with participatory monitoring in north-eastern Brazil: The case of AS-PTA's Projeto Paraiba. In: Estrella et al. (editors). Learning from Change: Issues and experience in participatory monitoring and evaluation. London: IT Publications
http://www.dfid.gov.uk/r4d/searchresearchdatabase.asp?ProjectID=1341

Spencer, Dunstan, Stein Bie, Ursula Blackshaw and Anne Thomson 2005, Evaluation of DFID's Renewable Natural Resources Research Strategy1995-2005. DFID EVSUM EV659
http://www.oecd.org/dataoecd/7/7/35242503.pdf

Spielman, D. and R. Birner. 2008. "How Innovative Is Your Agriculture? Using Innovation Indicators and Benchmarks to Strengthen National and Agricultural Innovation Systems." ARD Working Paper #41. Washington, DC: World Bank.

Swiss Agency for Development and Cooperation n/d Report on the Effectiveness of Swiss development cooperation in the agricultural sector in 2010 Technical analyses.
http://www.deza.admin.ch/en/Home/Effectiveness/Reports_on_effectiveness_and_annual_reports

Swiss Agency for Development and Cooperation 2010. Evaluation of SDC's Research Related Activities, Bern, March 2010 https://ext.d-nsbp
p.admin.ch/NSBExterneStudien/19/attachment/fr/70.pdf

UNEG 2005, Norms for Evaluation in the UN system. New York: United Nations Evaluation Group.
http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=21

Van den Berg, H. 2004. *IPM FFS: A synthesis of 25 impact evaluations. Report prepared for the Global IPM Facility*. Wageningen: The Netherlands.

Walker, T., M. Maredia, T. Kelley, R. La Rovere, D. Templeton, G. Thiele, and B. Douthwaite (2008) Strategic Guidance for *Ex Post* Impact Assessment of Agricultural Research. Report

prepared for the Standing Panel on Impact Assessment, CGIAR Science Council. Science Council Secretariat: Rome, Italy ftp://ftp.fao.org/docrep/fao/011/i0276e/i0276e.pdf

Weiss, C. H. 1995. Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families", in J. Connell, A. Kubisch, L. B. Schorr, and C. H. Weiss (eds) New Approaches to Evaluating Community Initiatives, New York: Aspen Institute.

White, H. 2009a. *Some Reflections on Current Debates in Impact Evaluation.* International Initiative for Impact Evaluation, Working paper 1. New Delhi: International Initiative for Impact Evaluation (3ie). http://www.3ieimpact.org/admin/pdfs_papers/11.pdf (accessed 6 June 2011).

White, H. 2009b. *Theory-Based Impact Evaluation: Principles and Practice.* International Initiative for Impact Evaluation, Working paper 3. New Delhi: International Initiative for Impact Evaluation (3ie). http://www.3ieimpact.org/admin/pdfs_papers/51.pdf (accessed 6 June 2011).

World Bank 2006 *Enhancing Agricultural Innovation: How to Go Beyond the Strengthening of Research Systems.* Washington: The International Bank for Reconstruction and Development/The World Bank.

http://siteresources.worldbank.org/INTARD/Resources/Enhancing_Ag_Innovation.pdf (accessed 6 June 2011).

Appendix 1 - Request for Information

Natural Resources Institute

12th July, 2011

Formal request for information for EIARD commissioned study on donor approaches to impact assessment of Agricultural Research for Development.

Dear Sir/Madam,

We have been commissioned to conduct a study for the European Initiative for Agricultural Research for Development (EIARD), the aim of which is to 'review and compare the policies and practices of different EIARD members in impact assessment (IA) to increase relevance, uptake and coordination of efforts by and for EIARD members, stakeholders and policy-makers'.

In order to conduct this study it is important that we gather information from all of the EIARD member organisations, firstly, on their approaches to understanding and measuring results from investments in ARD, and secondly that we gather examples of evaluation studies and impact assessments of ARD projects and programmes they have commissioned.

We, therefore, politely request that you fill in the attached table. We would also be grateful if you could share with us any relevant documents (such as policy documents outlining current ARD policy, impact assessments of agricultural programmes or evaluations of results of agriculture investments) or the URLs where these could be obtained.

We would be extremely grateful if you could you send this information to us as soon as possible, but by the **25th of July, 2011 at the latest** to Valerie Nelson, at the following email address: **v.j.nelson@gre.ac.uk**

Thank you very much,

Yours Sincerely,

Adrienne Martin (Director of Programme Development) &
Valerie Nelson (Livelihoods and Institutions Group)
Natural Resources Institute,
University of Greenwich,
Email: a.m.martin@gre.ac.uk
Tel: 01634-883055

Gilles Saint-Martin (CIRAD, FR)
Regional Director for South-East Insular Asia
Plaza Bisnis Kemang - Jl. Kemang Raya no. 2 - Jakarta 12730 - Indonesia
Tel: +62-21 7199067/4601. Fax: + 62-21 7179 3304
gilles.saint-martin@cirad.fr

Name of Donor	**Overall funding on Impact Assessment for ARD in last 5 years**	**Impact Assessments of ARD programmes and projects commissioned** (Please specify country, year, objective?, topic, partners/ collaborations) **Please attach report if possible, if not – provide URL*.	**Main Impact Assessment Methodologies used**	**Who were Impact Assessment Studies Conducted by?** (Please specify if internal evaluation or external commission consultants/ organisation /academic institutions?	**Ways in which the Impact Assessment has been used INTERNALLY** (Please specify how internal policies, and practices have been influenced)	**Ways in which the Impact Assessment has been used EXTERNALLY** (policy impacts, academic, practice)

Appendix 2 - Projects selected for analysis

PROJECT TITLE	No.	EIARD COUNTRY	SOURCE
Impact of Irrigation Development on Rural Poverty and the Environment	1	Austria	INFOSYS
Sub-Saharan Africa Challenge Programme (SSA CP)	2		INFOSYS
Reaping the Benefits: Assessing the impact and facilitating the uptake of resource-conserving technologies in the rice-wheat systems of the Indo-Gangetic Plain	3		INFOSYS
Participatory monitoring and output assessment of rural regeneration and sustainable agriculture in Brazil.	4	UK	DFID R4D
Optimising institutional arrangements for demand-driven post-harvest research, delivery, uptake and impact on the livelihoods of the poor, through public and private sector partnerships	5	UK	DFID R4D
Using livestock to improve livelihoods of landless and refugee-affected livestock keepers in Bangladesh and Nepal	6	UK	DFID R4D
Review of rice-wheat consortium for the Indo-Gangetic plains	7	UK	DFID R4D
Promotion and dissemination of integrated pest and soil fertility management strategies to combat striga, stemborers and declining soil fertility in the Lake Victoria basin	8	UK	DFID R4D
Impact assessment of community animal health services - a literature review	9	UK	DFID R4D
Promotion and impact assessment of tomato leaf curl virus disease resistant tomatoes: phase III of sustainable management and molecular characterisation of Bemisia tabaci and tomato leaf curl virus (ToLCV) on tomato in India	10	UK	DFID R4D
Access, Adoption and Diffusion: The Long-term Impacts of Improved Vegetable and Fish Technologies in Bangladesh	11		3IE
The Impact of Farmer-Field-Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes	12		3IE
The National Agriculture and Livestock Extension Programme (NALEP) Phase I Impact Assessment	13		SIDA
Conservation Agriculture in AFRICA: Analysing and Foreseeing its Impact - Comprehending its Adoption / 69 Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands	14	EC	3IE
Upland Rice Varieties for India Rainfed Agriculture Impact Study No. 1 Monitoring Impact Assessment and Learning Component (MIL) 2.2 of the Research into Use Programme (RiU)	15	UK	RIU
Rainfed Rabi Cropping in Rice-fallows of Nepal: Rainfed Agriculture Impact Assessment Study No. 4 Monitoring Impact Assessment and Learning Component (MIL) of the Research into Use Programme (RiU)	16	UK	RIU
Institutional innovations in India's crop improvement system: Rainfed Agriculture Impact Assessment Study No. 5	17	UK	RIU
The innovation trajectory of sleeping sickness control in Uganda: Research knowledge in its context	18	UK	RIU
Global 2007 - Evaluation of EC contribution to CGIAR	19	EC	EC
Global 2011 - Practical Application of CGIAR Research Results by Smallholder Farmers	20	EC	EC
Evaluation de la coopération belge au Rwanda	21	Belgium	Belgium
Evaluation de la coopération au développement bilatérale directe 1999-2005	22	Belgium	Belgium

Evaluation conjointe de la coopération de la Commission européenne, de la Belgique, du Danemark, de la France et du Luxembourg avec le Niger 2000-2008	23	Belgium	Belgium
"A cheval donné, on ne regarde pas les dents". Les mécanismes et les impacts de l'aide vus par des praticiens nigériens	24	Belgium	Belgium
Evaluation des évaluations 2000-2004	25	Belgium	Belgium
5 Year Ex-Post Impact Study POSTCOSECHA Programme Central America	26	Switzerland	SDC
Research to Impact: Case Studies for Natural Resource Management for Irrigated Rice in Asia	27	Switzerland	SDC
Water Saving in Rice Production-Dissemination, Adoption and Short Term Impacts of Alternate Wetting and Drying (AWD) in Bangladesh	28	Switzerland	SDC
Evaluation of SDC's Research Related Activities	29	Switzerland	SDC
Evaluation of SDC's Research Related Activities, The Policy Practice Limited, 2010		Switzerland	
SDC's contribution towards biodiversity: impact in the Andean region, Le Groupe-conseil Baastel, 2009	30	Switzerland	SDC
Impact assessment of push-pull technology developed and promoted by icipe and partners in Eastern Africa	31	Switzerland	SDC
Report on Effectiveness Swiss development cooperation in the agricultural sector 2010.Technical analyses: FAKT GmbH, Stuttgart Concept & Text: pluswert GmbH, Basel/Chur	32	Switzerland	SDC
Socio-cultural factors influencing adoption of ecologically based rodent pest management	33		
Impact assessment of increasing rice productivity in South & SE Sulawesi: water savings via AWD in the Philippines, monga mitigation from direct-seeded rice in Bangladesh	34		
Achieving Development Impact and Environmental Enhancement through Adoption of Balanced Nutrient Management Systems by Farmers in the West-African Savanna (BNMS II)	35	Belgium	**info sys**
Assessing the farm-level impact of genetically improved farmed tilapia (GIFT) in China and the Philippines	36	Germany ???? Or UK	**info sys**
Assessing the impact of the banana bacterial wilt, Xanthomonas campestris pv. musacearum on household livelihoods in East Africa	37	United Kingdom	**info sys**
Determination of high-potential aquaculture development areas and impact in Africa and Asia	38	Germany	**info sys**
Economic impact assessment as a decision-making tool for resource allocation in horticultural research in East Africa	39		**info sys**
Economic impact assessment of biological control of the diamond back moth in crucifers in East Africa	40		**info sys**
Field studies on the development and impact of drug resistance of animal trypanosomes in market-oriented production systems in the southern Guinea zone of West Africa	41	Germany	**info sys**

Impact of agricultural market reforms on smallholder farmers in Benin and Malawi	42	Germany	info sys
Impact of Consumption of the Crops Grown under Arsenic Rich Irrigated Water on Public Health in Nepal	43	Austria	info sys
Impact of production and marketing of freshwater aquatic products on rural livelihoods	44	United Kingdom	info sys
The impact of aquatic animal health strategies on the livelihoods of poor people in Asia	45	United Kingdom	info sys
Monitoring of developmental impact of 'Development and Promotion of African Rattans'	46	United Kingdom	info sys
The Impact of Organic Agriculture in Uganda: Improving Livelihoods through Sustainable Natural Resource Management and Market Linkages	47	Austria	info sys
The impact of organic cotton cultivation on the livelihood of Indian smallholders	48	Switzerland	info sys
Understanding rural livelihood systems for rice research prioritization and impact assessment	49	France	info sys
Understanding the impact of OECD cotton subsidies on developing countries and poor people in those countries	50		info sys
Assessing the contribution of diversified Musa genetic resources to poverty reduction, environmental sustainability and gender equality in rural communities	51	Germany	info sys
Monitoring and Evaluation: Pathways for Change - A Summary of Monitoring & Evaluation Experience from the Renewable Natural Resources Research Strategy (RNRRS) by Susanne Turrall and Kath Pasteur	52	UK	R4D
Learning from the Renewable Natural Resources Research Strategy Pathways for change: monitoring and evaluation	53		R4D
Synthesis of Monitoring and Evaluation Experience in the Renewable Natural Resources Research Strategy (RNRRS)	54	UK	R4D
EVALUATION OF DFID'S RENEWABLE NATURAL RESOURCES RESEARCH STRATEGY (RNRRS) 1995-2005Dunstan Spencer, Stein Bie, Ursula Blackshaw and Anne Thomson with ten subject matter specialists	55	UK	R4D
Synthesis report RIU: Monitoring & Evaluation of a Public-Private Partnership Stamp out Sleeping Sickness Case Study Catherine Butcher With contributions from J.F. Morton and Alexandra Shaw	56	UK	RIU
RESEARCH INTO USE PROGRAMME, IMPACT EVALUATION COMPONENT, WHAT HAVE WE LEARNT SO FAR? An end of contract report from the component task manager, Sheelagh O'Reilly, October 2009.	57		RIU
Results-Based Approach in Finnish Development Cooperation	58	Finland	
The Sustainability Dimension in Addressing Poverty Reduction: Synthesis of Evaluations	59	Finland	
Meta-Analysis of development evaluations 2007-8 Ministry of Foreign Affairs	60	Finland	
Caribbean and Banana Tariffs study	61		info sys

Evaluation of ecological and economic sustainability of breeding strategies in pastoral systems: The case of Ankole cattle in Uganda	62		**info sys**
Process and partnership for pro-poor policy change	63		**info sys**
CIMMYT - New Seed Initiative for Maize in Southern Africa (former SADLF)	64		**info sys**
Community-based livestock improvement: - a case study of farm-Africa's goat improvement project in Meru, Kenya:	65	UK	RIU
Monitoring, Impact Assessment and Learning for Research Into Use – Nigeria	66	UK	RIU
Dissemination of Improved Bambara Processing Technologies Through a New Coalition Arrangement to Enhance Rural Livelihoods in Northern Ghana	67	UK	RIU
The Impact of Improved Maize Varieties on Poverty in Mexico: A Propensity Score-matching Approach	68		3IE
Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands	69		
Technology adoption, productivity and specialization of Uruguayan breeders: evidence from an impact evaluation	70		3IE

	✓	TL
R4D	9	11
other	19	20
RIU	5	9
Info sys	12	24
3ie	5	6
	50	**70**

50 projects selected for further analysis

	Projects not selected
	Projects combined with another entry for the same project

The characteristics of ARD (from: EIARD Strategy 2009-13, Annex 1)

ARD is intrinsically:

- ***fundamental & applied*** – dealing with upstream and problems solving research;
- ***comprehensive*** – dealing potentially with research objects in any field and at any relevan scale, thus encompassing a wide range of scientific disciplines (from molecular biology o genetics to agroecology; economics, political and social sciences or modelling);
- ***multi-stakeholder*** – because concerned people are many and face a variety of often ill known specific situations, thus requiring iterative and inter-active loops of participator diagnosis-to research-product processes that include all players and activities of the loca innovation systems.
- ***international*** – because carried out in and/or for developing & emerging econom countries, and in most cases with Southern ARD partners and International Agricultura Research Centres (in particular CGIAR Centres).
- ***global*** – as similar problems are widely shared among countries and as local interaction with world problems result from globalizations of all kinds;
- ***multiple policy*** purposed – because it contributes to various and different policies Science & Research, Agriculture, Foreign Affairs & Development, Environment, Trade & Economy, Health policies, to mention the most important ones.
- The concept of ARD has evolved considerably over the last decade from research whic focused directly on reducing hunger to the wider issues concerned with poverty alleviation, and is now beginning to address the challenge of sustainable development ***fo all*** within the concepts of ***"One World"*** and ***"Global Changes"***.
- ARD is now expected to broaden its agenda towards challenges of mutual interest o developing, emerging and industrialised countries

IMPACT ASSESSMENT POLICIES AND PRACTICES OF EIARD MEMBERS | **STUDY**

Adrienne Martin and Valerie Nelson

AGRINATURA is a new alliance formed by 35 European universities and research organisations working in agricultural research, education, training and capacity strengthening for development.
AGRINATURA members are involved in a broad range of issues related to agricultural research and education for development contributing through their expertise and experience.
AGRINATURA focuses on initiatives that open up new opportunities for farmers to enhance food security and improve the agro-food sector in general, whilst reducing the negative impact of agricultural activities on the environment.
Thanks to AGRINATURA's unparalleled access to major research institutions and universities in Europe and the rest of the world, it is able to nurture scientific excellence through training and exchanges and further sustainable development in agriculture through joint research and education programmes and projects.

AGRINATURA formulates and implements research and education programmes and projects in developing and emerging economy countries on every continent.

At the practical level, AGRINATURA partners interact with a single office (the management unit) that:
- can widely inform the European ARD community of partnerships opportunities;
- can directly enter partnerships and consortia that can respond to the Agrinatura objectives;
- can mobilise necessary experts from 31 research, training and development organisations to work almost anywhere.

AGRINATURA assets are:
- global coverage of key issues in agricultural research for development, focusing mainly on developing countries and countries with emerging economies;
- a broad spectrum of complementary expertise in disciplinary and interdisciplinary research and development which allows AGRINATURA to work at the interfaces;
- solid experience in integrative and participatory approaches at different scales;
- translation of development issues into a researchable agenda;
- inclusion of development projects into on-going research and education programmes;
- partnership which goes beyond the function of services provider; regular and continuous contacts with project partners in the field before, during and after operation of programmes;
- extensive experience in capacity development and scientific support for the formulation of international development policies, and the search for project funding thanks to its collaboration with and support for partner institutions and stakeholders.

For further information on AGRINATURA Association:

AGRINATURA Association Secretariat
Czech University of Life Sciences Prague
Kamýcká 129, 165 21 Prague 6 Czech Republic
Phone: +420 224 382 011
Fax: +420 224 382 012
secretariat@agrinatura.eu
http://www.agrinatura.eu/

For further information on AGRINATURA-EEIG:

AGRINATURA-EEIG Secretariat
42 rue Scheffer
F-75116 PARIS
FRANCE
Fax: +33.1.53.70.21.56
secretariat@agrinatura-eeig.eu
http://www.agrinatura.eu/