

Working Paper 281

**Making a difference:
M&E of policy research**

Ingie Hovland

July 2007

Overseas Development Institute
111 Westminster Bridge Road
London
SE1 7JD
UK

ISBN 978 0 85003 848 4

© Overseas Development Institute 2007

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publishers.

Contents

Acknowledgements	v
Acronyms and abbreviations	v
Executive Summary	vii
1 Introduction: The non-academic impact of research	1
1.1 Target audience	1
1.2 Beyond citations	1
1.3 Talking about impact: a few definitions	1
1.4 A short note on levels: projects, programmes, and institutions	3
1.5 Overview of the paper	3
2 Lessons from current experience: How to evaluate five key performance areas of policy research projects, programmes, and institutions	4
2.1 Performance Area I: Evaluating strategy and direction	4
2.2 Performance Area II: Evaluating management	16
2.3 Performance Area III: Evaluating outputs	20
2.4 Performance Area IV: Evaluating uptake	26
2.5 Performance Area V: Evaluating outcomes and impacts	27
3 Institutional evaluations: Additional concerns	36
4 Conclusion: Designing an M&E plan for a policy research project, programme or institution	38
4.1 Adding value	38
4.2 Best practice checklist for policy research projects	39
4.3 Best practice checklist for policy research programmes	39
4.4 Best practice checklist for policy research institutions	40
4.5 When you are asked to evaluate someone else’s project, programme or institution	43
References	44
Appendix 1. Examples of institutional evaluations of policy research institutes and think tanks	48

Boxes and figures

Box 1. CGIAR, including SPIA, ILAC, LTI, CIAT, IFPRI and ACIAR	13
Box 2. DFID	16
Box 3. ECDPM	18
Box 4. IDRC	29
Box 5. ODI	35
Figure 1. Example of a Problem Tree	5
Figure 2. Stakeholder Analysis Grid	5
Figure 3. Outline of a Logframe	6
Figure 4. Example of a SNA map	7
Figure 5. The process of constructing Impact Pathways and Narrative	9
Figure 6. Example of Problem Tree from the CPWF: Water and People in Catchments	10
Figure 6 (continued). Example of Objective Tree from the CPWF: Water and People in Catchments	11
Figure 7. Example of a Gantt chart	12
Figure 8. In the process of constructing a Network Map for one of the CPWF projects	13
Figure 9. Example of an Outputs x Impacts matrix	15
Figure 10. Example of an Outputs x Stakeholder matrix	15
Figure 11. The Appreciative Inquiry '5-D' model	19
Figure 12. After Action Review questions	25
Figure 13. The three stages of Outcome Mapping	28
Figure 14. Example of ROA Output	30
Figure 15. Example of an Actor x Actor matrix	34
Figure 16. The CAF model	36
Figure 17. The POET model	37
Figure 18. IISD Policy Influence Report Card	54

Acknowledgements

The author would like to thank Fred Carden and James McGann for helpful peer review comments. Thanks also to colleagues at ODI for useful discussions, especially Simon Maxwell and John Young.

For further information on the work of the Research and Policy in Development Group (RAPID) at ODI, or for information on ODI's work on M&E in general, please contact John Young, Director of Programmes for RAPID, at j.young@odi.org.uk.

Ingie Hovland is a Research Officer in the Research and Policy in Development Group (RAPID) at ODI. She works on communication, M&E and the research-policy interface.

Acronyms and abbreviations

AAR	After Action Review
ACIAR	Australian Centre for International Agricultural Research
ACP	African, Caribbean and Pacific Group of States
AI	Appreciative Inquiry
CAF	Common Assessment Framework
CEO	Chief Executive Officer
CGD	Centre for Global Development
CGIAR	Consultative Group on International Agricultural Research
CHSRF	Canadian Health Services Research Foundation
CIAT	International Centre for Tropical Agriculture
CPHP	DFID's Crop Post-Harvest Research Programme
CPWF	CGIAR's Challenge Programme on Water and Food
CRD	DFID's Central Research Department
CSO	Civil Society Organisation
CSPP	ODI's Civil Society Partnership Programme
CTA	Technical Centre for Agricultural and Rural Cooperation ACP-EU
DFID	The UK Department for International Development
DRC	Development Research Centre
ECDPM	European Centre for Development Policy Management
EngKaR	DFID's Engineering Knowledge and Research Programme
EU	European Union
FAO	Food and Agriculture Organisation of the UN
HTML	Hypertext Mark-up Language
ICT	Information and Communication Technology
IDRC	International Development Research Centre
IDS	Institute for Development Studies
IFPRI	International Food Policy Research Institute
IISD	International Institute for Sustainable Development
ILAC	CGIAR's Institutional Learning and Change Initiative
ILRI	International Livestock Research Institute
INTRAC	International NGO Training and Research Centre
LTI	Learning to Innovate Group
M&A	Monitoring and Assessment
M&E	Monitoring and Evaluation
MOV	Means of Verification
MSC	Most Significant Change
NGO	Non-government organisation

NSI	National Systems of Innovation
ODI	Overseas Development Institute
OECD	Organisation for Economic Cooperation and Development
OECD-DAC	OECD Development Assistance Committee
OPM	Oxford Policy Management
PARC	Performance Assessment Resource Centre
POET	Participatory Organisational Evaluation Tool
PPPPC	Process and Partnership for Pro-poor Policy Change
PRSP	Poverty Reduction Strategy Paper
QAA	The UK Quality Assurance Agency for Higher Education
RAE	Research Assessment Exercise
RAPID	ODI's Research and Policy in Development Group
RBM	Results-Based Management
R&D	Research and Development
RNRRS	DFID's Renewable Natural Resources Research Strategy
ROA	RAPID Outcome Assessment
SC	Significant Change
SCI	Science Citation Index
SDP	Smallholder Dairy Project
SIDA	Swedish International Development Agency
SNA	Social Network Analysis
SOAS	School of Oriental and African Studies
SPIA	CGIAR's Standing Panel on Impact Assessment
SWOT	Strengths, Weaknesses, Opportunities, Threats
TQM	Total Quality Management
UK	United Kingdom
UNDP	United Nations Development Programme
WGIE	IFPRI's Working Group on Impact Evaluation

Executive Summary

This paper aims to advance understanding on how to monitor and evaluate policy research, i.e. research that is undertaken in order to inform and influence public policy. Policy is defined very broadly to encompass both policy decisions and processes, including implementation.

Conventional academic research is usually evaluated using two approaches: academic peer review, and number of citations in peer-reviewed publications. For policy research programmes, these evaluation tools have proven too limited. They are not well suited to capture some of the broader aims of policy research, such as policy impact, changes in behaviour, or building of relationships. In short, policy research programmes need new monitoring and evaluation (M&E) approaches in order to know whether they are making a difference, not only in the academic world but also in the world outside academia.

The paper is written with research programmes and institutions in mind, rather than individual researchers. It presents examples and approaches on how to do M&E of policy research from the current experience of a range of research institutes, think tanks and funding bodies. The approaches have been divided into the following five key performance areas: (i) Strategy and direction; (ii) Management; (iii) Outputs; (iv) Uptake; and (v) Outcomes and impacts. Research programmes or institutes may wish to focus on only one of these areas, or may combine approaches across the areas to form a more comprehensive M&E plan.

This paper has five sections. Section 1 is a short introduction. Section 2 provides a survey of a range of possible new M&E approaches taken from the current experience of policy research projects, programmes and institutions. These are:

- **Performance Area I – Evaluating strategy and direction:** Logframes; Social Network Analysis; Impact Pathways; Modular Matrices
- **Performance Area II – Evaluating management:** ‘Fit for Purpose’ Reviews; ‘Lighter Touch’ Quality Audits; Horizontal Evaluation; Appreciative Inquiry
- **Performance Area III – Evaluating outputs:** Evaluating academic articles and research reports; Evaluating policy and briefing papers; Evaluating websites; Evaluating networks; After Action Reviews
- **Performance Area IV – Evaluating uptake:** Impact Logs; New Areas for Citation Analysis; User Surveys
- **Performance Area V – Evaluating outcomes and impacts:** Outcome Mapping; RAPID Outcome Assessment; Most Significant Change; Innovation Histories; Episode Studies

Section 2 also presents notes on institutions that have begun developing new models in the area of M&E of policy research. These are CGIAR (including SPIA, ILAC, LTI, CIAT, IFPRI and ACIAR), DFID, ECDPM, IDRC and ODI. Further details on institutional evaluations are given in Appendix 1.

Section 3 then highlights a few additional concerns to bear in mind when evaluating entire institutions (rather than individual projects or programmes), and Section 4 concludes by presenting best practice checklists on how to design an M&E approach for a policy research project, programme, or institution.

1 Introduction: The non-academic impact of research

1.1 Target audience

This paper aims to advance understanding on how to monitor and evaluate policy research, i.e. research that is undertaken in order to inform and influence public policy. Policy is defined very broadly to encompass both policy decisions and processes, including implementation.

The paper is written with research programmes and institutions in mind, rather than individual researchers. It presents examples and approaches on how to do monitoring and evaluation (M&E) of policy research from the current experience of a range of research institutes, think tanks and funding bodies. The approaches have been divided into the following 'key performance areas' in this paper: (i) Strategy and direction; (ii) Management; (iii) Outputs; (iv) Uptake; and (v) Outcomes and impacts. Research programmes or institutes may wish to focus on only one of these areas, or may combine approaches across the areas to form a more comprehensive M&E plan.

1.2 Beyond citations

The establishment of the Science Citation Index (SCI) in the 1960s was a major breakthrough in the field of research evaluation. Since then, conventional academic research has usually been evaluated using two approaches: academic peer review, and number of citations in peer-reviewed publications. Over the past few years, however, it has become increasingly clear that in the world of policy research, the mechanisms of academic peer review and conventional citation counting are starting to prove too limited.

They are not well suited to capture some of the broader aims of policy research, such as policy impact, changes in behaviour, or building of relationships. Neither are they well suited to evaluate some of the outputs that policy research programmes consider important, such as policy briefing papers, a website, public meetings, one-on-one meetings, coalitions and networks. Lastly, they do not take into account the range of stakeholders that policy research programmes often wish to communicate with, such as policymakers, bureaucrats, donors, businesses, civil society organisations, the media, or the public. Policy researchers may wish to be cited in academic journals, but they may just as well consider their research outputs to be successful if they are able to pose a development problem concisely enough for a policymaker to have time to read it and take notice; or if they get mentioned in a prominent newspaper article; or if their research has managed to create new links between a civil society organisation working on a particular topic and a private business in the same area.

In short, policy research programmes need new monitoring and evaluation (M&E) approaches that go beyond academic peer review and conventional citation counting, in order to know whether they are reaching their goals not only in the academic world but also in the world outside academia. This has also become a concern for development funders who commission research. They would like to know whether the research they fund makes a difference outside the academic community, and how and where to allocate research funds in order to contribute to overarching development goals. They would like to know about its *impact*.

1.3 Talking about impact: a few definitions

To some extent, the term 'impact' is already well-known within the development business, since it has been included in the OECD-DAC criteria for evaluating development assistance: relevance, efficiency, effectiveness, impact and sustainability. Policy researchers within the international development field are therefore familiar with the concept. When the idea of 'impact' is applied to policy research,

however, it quickly becomes clear that this raises different questions than when the idea of ‘impact’ is applied to development assistance initiatives. In order to measure the impact of policy research (as well as its relevance, efficiency, effectiveness and sustainability), we need processes and methods that adequately capture the complexities of research-policy linkages.

Firstly, indirect impacts, outcomes and changes are an important part of the non-academic impact of research. These impacts are hard to pin down. As Davies, Nutley and Walter put it:

Non-academic research impact is about identifying the influences of research findings on policy, managerial and professional practices, social behaviour or public discourse. Such impact may be *instrumental*, influencing changes in policy, practices and behaviour, or *conceptual*, changing people’s knowledge, understanding and attitudes towards social issues. (Davies, Nutley and Walter, 2005: 11, original emphasis)

Both instrumental and conceptual impacts of research are difficult to measure. When research has an *instrumental* impact on policy or practice, this often occurs in conjunction with a series of other events and relationships, and thus the relative contribution of the research to the outcome is not easily determined. This difficulty is enhanced even further when it comes to *conceptual* impacts, where research may have been converted into an anecdote, a catchphrase, or received wisdom. In these cases the research may have ‘percolated’ through various policy and practitioner networks, to great effect, but without being tagged as a specific piece of research (Weiss, 1977).

In order to get to grips with the concept of impact, it may be useful to look at a few key terms that are used in this area: outputs, outcomes and impacts. While these terms are used ubiquitously in the literature on evaluation, there are no universal definitions for them. Results-Based Management (RBM) uses the model *Inputs* › *Activities* › *Outputs* › *Outcomes* › *Impact*, where ‘outcomes’ refers to mid-term accomplishments, and ‘impacts’ refers to long-term results. The Impact Pathways model used by the CGIAR defines an ‘outcome’ as the external use, adoption, or influence of programme outputs, leading to ‘changes in knowledge, attributes, policies, research capacities, agricultural practices, productivity, sustainability or other factors required in order to achieve the intended impact’ (Douthwaite et al., 2006: 9), while ‘impacts’ are defined as any longer-range benefits. The Outcome Mapping approach developed by IDRC uses a model that focuses on ‘outcomes’, which refers to changes in behaviour, and ‘impacts’, which are seen as longer-term goals. In much of the Outcome Mapping literature, however, ‘outcomes’ and ‘impacts’ are simply used as synonyms (Earl, Carden and Smytulo, 2001).

Given the various working definitions that are in use, this paper will not enter into the debate over the exact distinction between e.g. outcomes and impacts. Rather, drawing on the definitions used above, it will take as its starting point the following broad working definitions:

- **Strategy and direction:** The basic plan that the research project/programme/institution is following in order to reach its intended goals.
- **Management:** The systems and processes that the project/programme/institution has in place in order to ensure that the overall strategy is carried out and that high-quality policy research is produced (e.g. systems of peer/user review, quality assurance, planning cycles, etc).
- **Outputs:** The tangible goods and services that a research project/programme/ institution produces (e.g. Working Papers, journal articles, policy briefs, website, meetings, events, networks, etc).
- **Uptake:** Direct responses to the research project/programme/institution (e.g. its research is mentioned in a government policy paper, on a range of websites, referred to in a newspaper article, etc).
- **Outcomes and impacts:** Changes in behaviour, knowledge, policies, capacities and/or practices that the research has contributed to, directly or indirectly (e.g. a change in government policy implementation, a change in working practices among NGO practitioners, a reduction of poverty in a certain area, strengthened livelihoods, strengthened civil society input into policy processes, etc).

1.4 A short note on levels: projects, programmes, and institutions

Section 2 of this paper presents various approaches that can be used to carry out M&E of policy research, drawn from the current experience of policy research projects, programmes, and institutions. To some extent, the difference in levels between projects, programmes and institutions needs to be taken into account when choosing an M&E approach. However, the M&E approaches that can be used for policy research projects will to some extent overlap with those that can be used for policy research *programmes* – for example, both projects and programmes will need to be planned in a way that takes into account key stakeholder groups, intended impacts, outputs, etc. Similarly there is some overlap between M&E approaches that can be used at programme level with those that can be used at an institutional level. In order to clarify where approaches overlap, Section 2 will indicate the level (project, programme, and/or institution) that is most appropriate for each approach.

1.5 Overview of the paper

This paper has five sections. Section 2 provides a survey of a range of possible M&E approaches taken from the current experience of policy research projects, programmes and institutions. Clearly, most of these approaches can be used across several of the key performance areas, but for the sake of simplicity they will be presented in the following order:

- **Performance Area I – Evaluating strategy and direction:** Logframes; Social Network Analysis; Impact Pathways; Modular Matrices
- **Performance Area II – Evaluating management:** ‘Fit for Purpose’ Reviews; ‘Lighter Touch’ Quality Audits; Horizontal Evaluation; Appreciative Inquiry
- **Performance Area III – Evaluating outputs:** Evaluating academic articles and research reports; Evaluating policy and briefing papers; Evaluating websites; Evaluating networks; After Action Reviews
- **Performance Area IV – Evaluating uptake:** Impact Logs; New Areas for Citation Analysis; User Surveys
- **Performance Area V – Evaluating outcomes and impacts:** Outcome Mapping; RAPID Outcome Assessment; Most Significant Change; Innovation Histories; Episode Studies

Section 2 also presents Boxes on institutions that have begun developing new models in the area of M&E of policy research. These are CGIAR (including SPIA, ILAC, LTI, CIAT, IFPRI and ACIAR), DFID, ECDPM, IDRC and ODI. Further details on institutional evaluations are given in Appendix 1.

Section 3 then highlights a few additional concerns to bear in mind when evaluating entire institutions (rather than individual projects or programmes), and Section 4 concludes by presenting best practice checklists on how to design an M&E approach for a policy research project, programme, or institution.

2 Lessons from current experience: How to evaluate five key performance areas of policy research projects, programmes, and institutions

2.1 Performance Area I: Evaluating strategy and direction

This section gives examples of approaches that have been used to monitor and evaluate to what extent a policy research project/programme/institution is being strategic, i.e. whether its basic plan and direction will lead to its intended goals.

2.1.1 Logframes

Level: Projects, Programmes, Institutions

Despite sustained criticism, logframes are still the most commonly recognised form of M&E within international development, and this has influenced the way policy research is evaluated too. Many donors (e.g. DFID) ask policy research programmes to draw up a logframe as the basis of their M&E. This section will briefly present a full logframe approach and suggestions on how its terminology may be used in a policy research project or programme, drawing on the definitions that have been chosen for the Civil Society Partnership Programme (CSPP) at ODI.

A Logframe matrix can be constructed on its own. However, a full Logical Framework Approach is meant to involve: (1) problem analysis (problem tree), (2) objectives tree, (3) objectives hierarchy, (4) stakeholder analysis, and (5) selecting a preferred strategy, including outputs and activities. The product of this analytical approach is (6) the matrix (the Logframe), which summarises what the project, programme or institution intends to do and how, what the key assumptions are, and how outputs and outcomes will be monitored and evaluated.

(1) **The Problem Tree:** Problem tree analysis is best carried out in a small focus group of about six to eight people using flip chart paper or similar. It is important that factors can be added as the conversation progresses. The first step is to discuss and agree the problem or issue to be analysed. The problem or issue is written in the centre of the flip chart and becomes the ‘trunk’ of the tree. This becomes the ‘focal problem’. Next, the group identify the causes of the focal problem – these become the roots below it – and then identify the consequences – which become the branches above it. The heart of the exercise is the discussion that is generated as factors are arranged and re-arranged. An example is given in Diagram 1 below, and another example is given in Diagram 6.

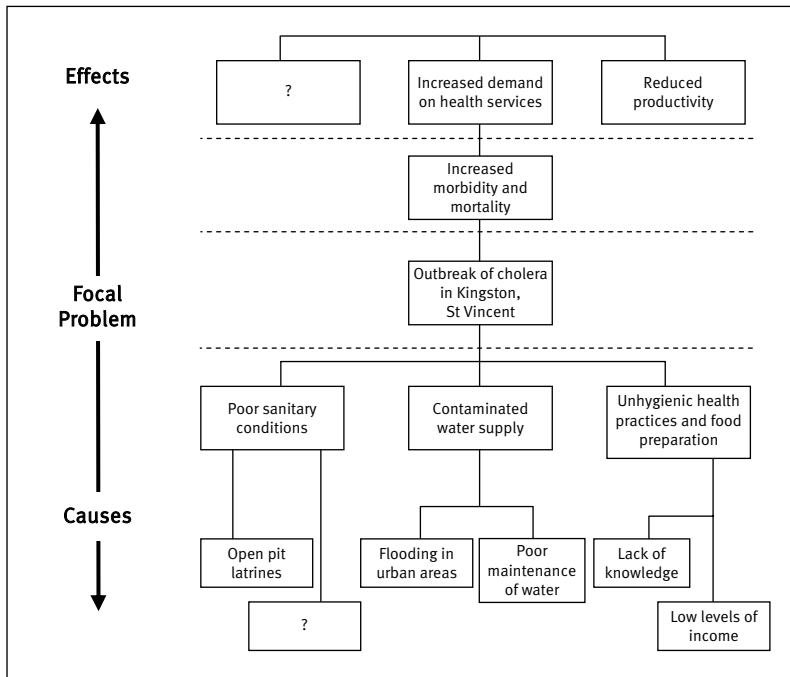
(2) **The Objectives Tree:** The Problem Tree can be converted into an Objectives Tree by rephrasing each of the problems into positive desirable outcomes – as if the problem had already been treated. In this way, root causes and consequences are turned into root solutions, and key project or influencing entry points are established. These objectives may well be worded as objectives for change. (See Figure 6.)

(3) **An Objectives Hierarchy:** The objectives that have been identified are turned into project or programme priorities. These may then feed into a Force Field Analysis which could provide a useful next step before the Stakeholder Analysis or Social Network Analysis (for more information on Force Field Analysis, see e.g. Hovland, 2005).

(4) **Stakeholder Analysis:** A stakeholder is a person who has something to gain or lose through the outcomes of the policy research project or programme. Stakeholder Analysis groups stakeholders into a four-field grid depending on their power and interest.

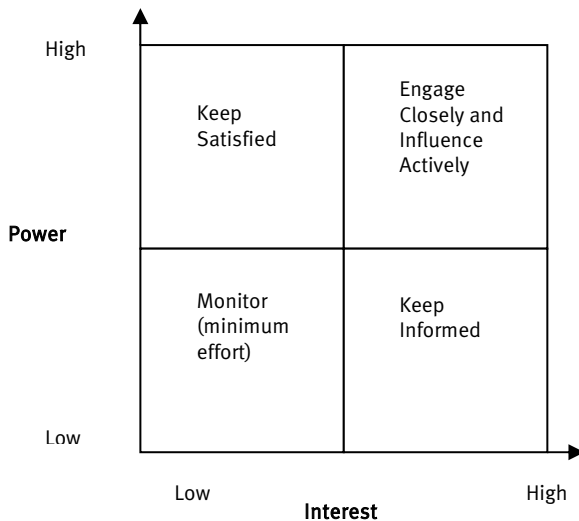
Figure 1. Example of a Problem Tree

The immediate 'issue' is put in the middle; the bottom half of the 'tree' then represents possible root causes and the top half represents consequences.



Source: DFID (n.d.)

Figure 2. Stakeholder Analysis Grid



More fine-tuned and fluid versions of Stakeholder Analysis can be carried out using Social Network Analysis (described in the section below).

(5) **Strategy:** Discuss and design a strategy for the project or programme (or institution), with a detailed list of outputs and activities, bearing in mind the Problem Tree, Objectives Tree, Objectives Hierarchy, and Stakeholder Analysis.

(6) **Draw up the Logframe:** The figure below gives an outline of a logframe.

Figure 3. Outline of a Logframe

	Narrative Summary	Indicators	Means of Verification (M&E)	Assumptions
Goal				-
Purpose				(Purpose to Goal)
Output (1)				(Output to Purpose)
Output (2)				
Output (3)				
Output (4)				
Activities		Inputs		(Activities to Output)

Logframe terminology can be defined in different ways. Some suggested definitions (adapted from ODI's Civil Society Partnerships Programme) are:

Goal: The overall goal that the project or programme is making a contribution towards.

Purpose: Observable changes in the behaviour of real people (e.g. our partner CSOs make more use of evidence). The project or programme should influence and contribute to the Purpose, but cannot be held directly accountable for its achievement (due to the number of other determining factors involved). The Purpose(s) should be designed in relation to the intended Outcomes and Impacts of the project or programme (cf. Section 2.5 below on Outcomes and Impacts).

Outputs: Tangible goods and services that the programme will produce and that others can use (e.g. Toolkit, Working Paper, Network website) in order to achieve the Purpose. The programme is directly responsible for the Outputs and can be held accountable for these.

Activities: What the project or programme is doing in order to produce the Outputs (e.g. We will review the literature, or We will host a workshop, or We will build and strengthen relationships).

Inputs: The financial, technical and human resources required in order to do the Activities. They are listed in the same row as Activities (instead of Indicators).

Narrative summary: Descriptive statement on Purpose, Outputs, and Activities.

Indicators (or Objectively Verifiable Indicators): The measurable changes that need to happen in order to achieve the Purpose and Outputs. Indicators may specify quantity, quality, and timing. Measurement may relate to:

- process (e.g. CSO sets up a process of internal review in the period xx – xx);
- product (e.g. CSO has adopted a written strategy by xx); and/or
- impact (e.g. X number of CSO staff report closer relationships with policymakers in current year than in previous year).

Means of Verification (MOVs): These are the M&E tools that are used to find out whether the measurable changes have taken place (e.g. Formal or Informal Interviews / After Action Reviews / Surveys / Expert Peer Reviews / Impact Box / Outcome Mapping / Most Significant Change / etc).

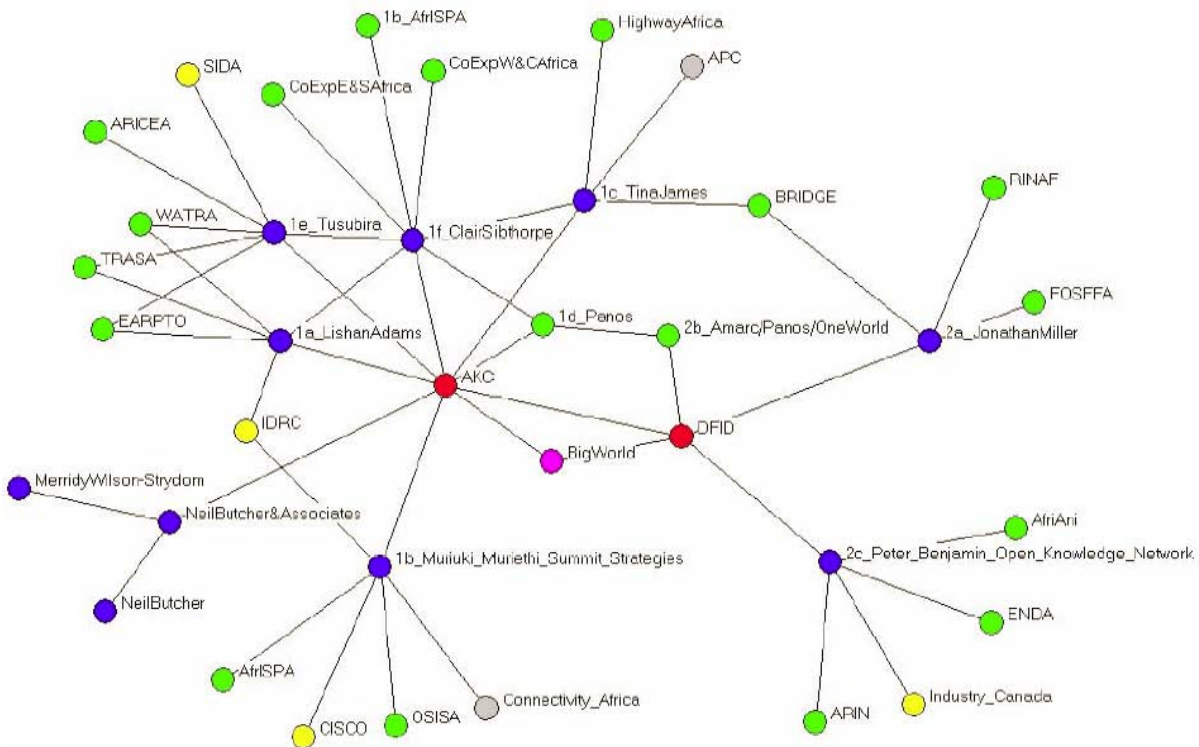
Assumptions: Other events or conditions that are necessary for Activities to lead to Output, Outputs to lead to Purpose, and Purpose to lead to Goal.

2.1.2 Social Network Analysis

Level: Projects, Programmes, Institutions

SNA is a process of learning about and understanding the (formal and informal) networks that operate in a given field. A range of methods can be used, including ethnography, participant observation, key informant interviews, semi-structured interviews, 'snowball' sampling, focus groups, and content analysis of the media (Schelhas and Cervený, 2002). The aim is to construct a 'map' of the linkages that exist between people in this field.

Figure 4. Example of a SNA map



Source: Davies (2003: 6)

The figure above gives an example of a SNA map. It has been constructed by Rick Davies to show both the existing and potential stakeholder linkages in an Africa ICT programme.

A rigorous and academic SNA should involve a trained social scientist. However, this level of analysis is often not necessary for practitioners and applied researchers who are only using SNA as a means to an end, namely more efficient collaboration with others (Schelhas and Cervený, 2002). For this purpose, an exploratory map of existing and potential linkages between actors is often sufficient. Software can also be used to analyse and plot the maps that have been drawn, such as NetDraw (www.analytictech.com/downloadnd.htm), Inflow (www.orgnet.com), or Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>).

Some key questions for SNA (adapted from Schelhas and Cervený, 2002: 3–4) are:

- Who are the relevant groups and individuals involved in or affected by issues in this policy research field?
- Are there identifiable groups or sub-groups (e.g. based on location, profession, interests, values, race, ethnicity, class or gender)?
- What are the past and present relationships between them?

- Who trusts whom?
- Who and what groups have power, and what is the source of their power?
- Who are the formal and informal leaders in the field?
- How do people exchange information?
- Do networks change, e.g. are they seasonal, or do they vary around issues?
- What else is important in this particular field?

2.1.3 *Impact Pathways*

Level: Projects. Can be used at programme level, as done by CGIAR, if all projects in the programme follow the same approach.

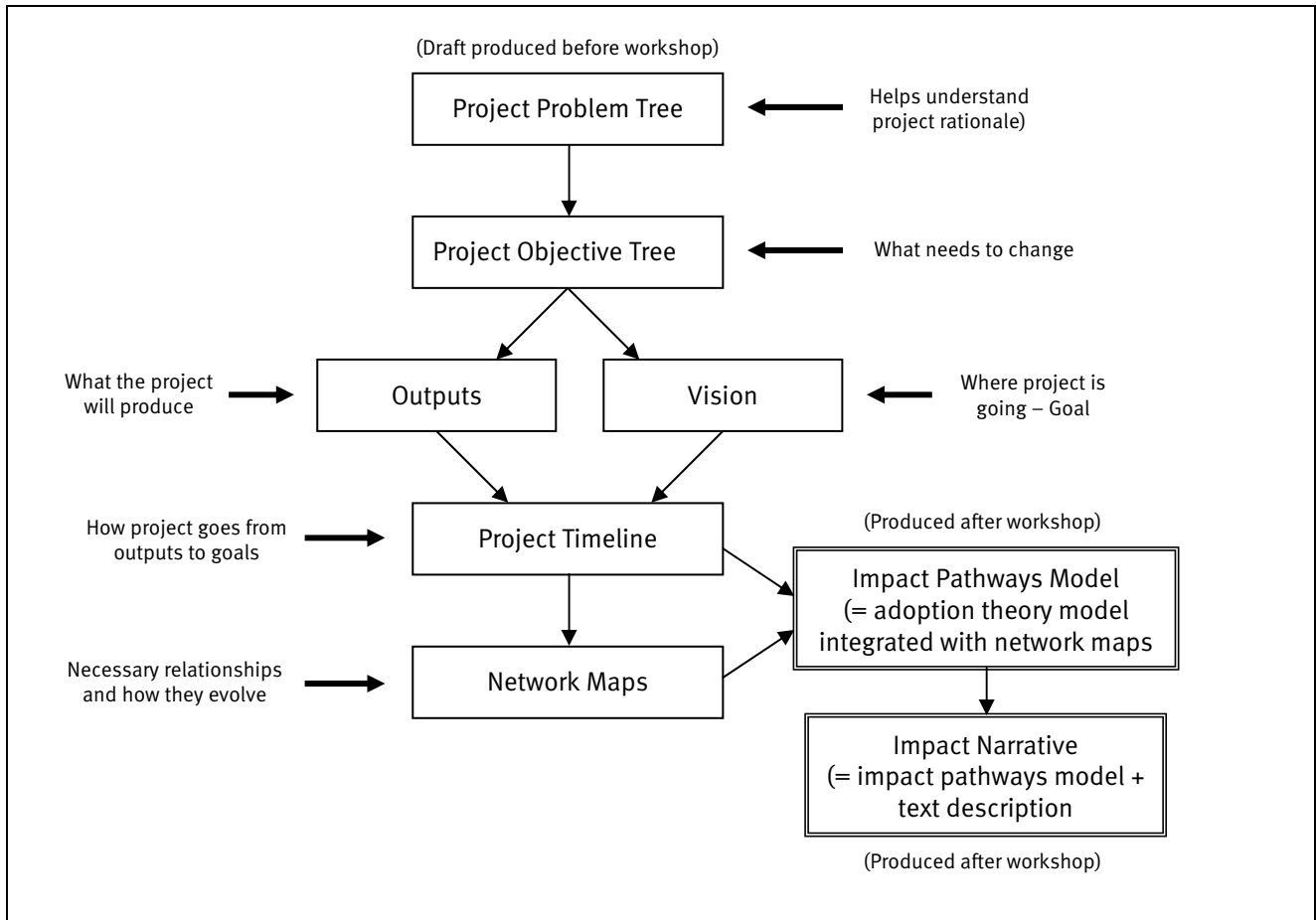
The Impact Pathways model draws on Social Network Analysis. It is being used by the Consultative Group on International Agricultural Research (CGIAR) in order to evaluate its Challenge Programme on Water and Food (CPWF), a research programme that consists of 51 research projects across five research themes, within nine river basins (the Limpopo, Nile, Yellow, São Francisco, Karkheh, Mekong, Nile, Volta, and the Andean system). In 2006 the number of participating institutions numbered 198, including CGIAR centres, research institutes, NGOs, community-based organisations, and national agricultural research and extension systems.

In short, Impact Pathway models ‘specify how networks of actors develop and use project or programme outputs to help generate chains of intermediate outcomes and eventual impact’ (Douthwaite, et al., 2006: 2). They do this through combining two core elements:

- a logical model, such as a logframe, which defines the causal chains from activities and outputs to outcomes and goals;
- a network model, such as Social Network Analysis, which shows the evolving relationships between programme organisations and other partners and stakeholders that are necessary to achieve impacts.

These elements will come together in, for example, a programme planning workshop, or a mid-term review workshop, which might produce the documents outlined in Figure 5.

Figure 5. The process of constructing Impact Pathways and Narrative



Source: Douthwaite, et al. (2006: 5)

An example of *Problem and Objective Trees* from a CPWF project are given in Figure 6 below. They show the link between logframe outputs (in the Objective Tree) and eventual impact.

Figure 6. Example of Problem Tree from the CPWF: water and people in catchments

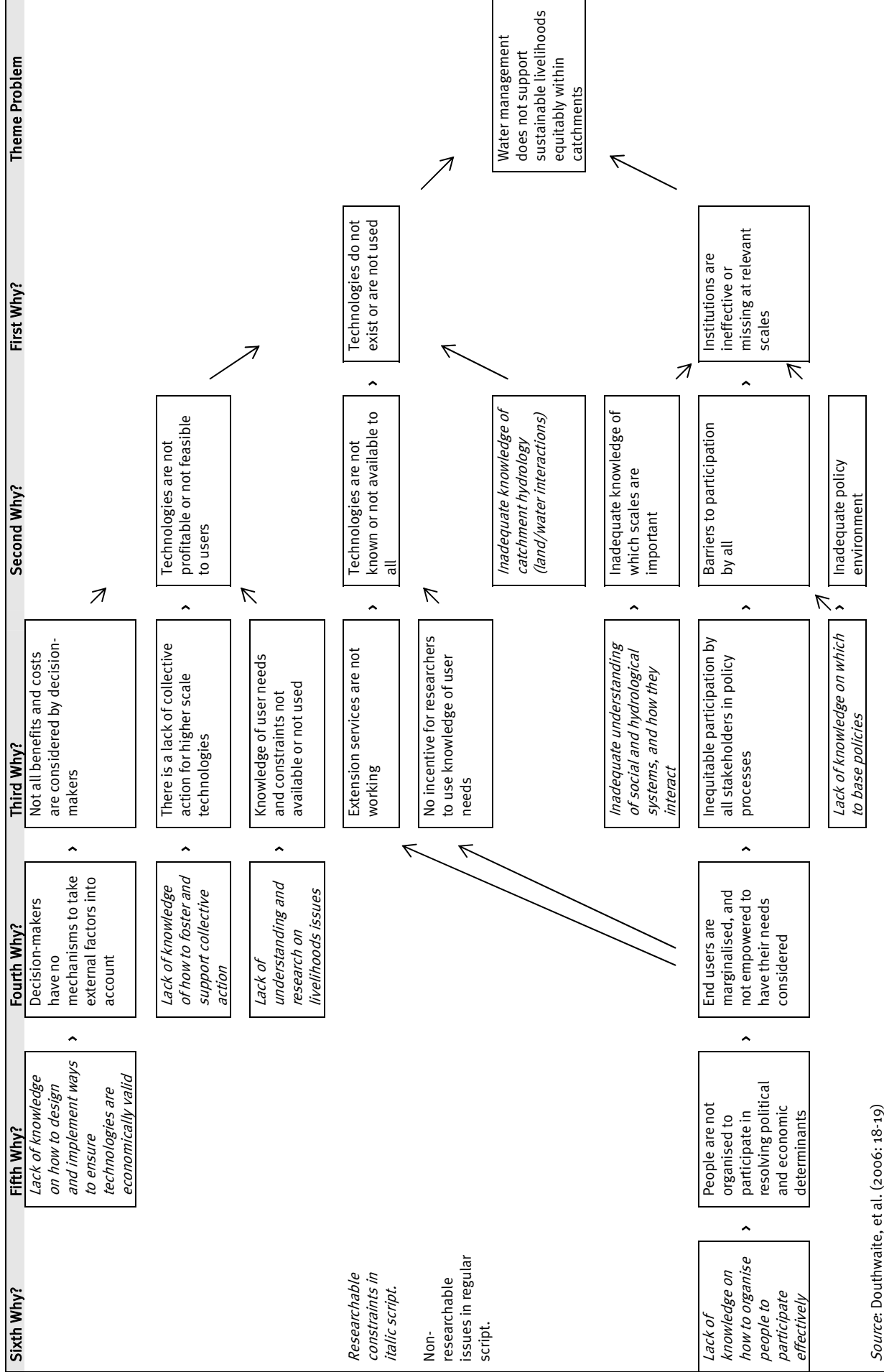
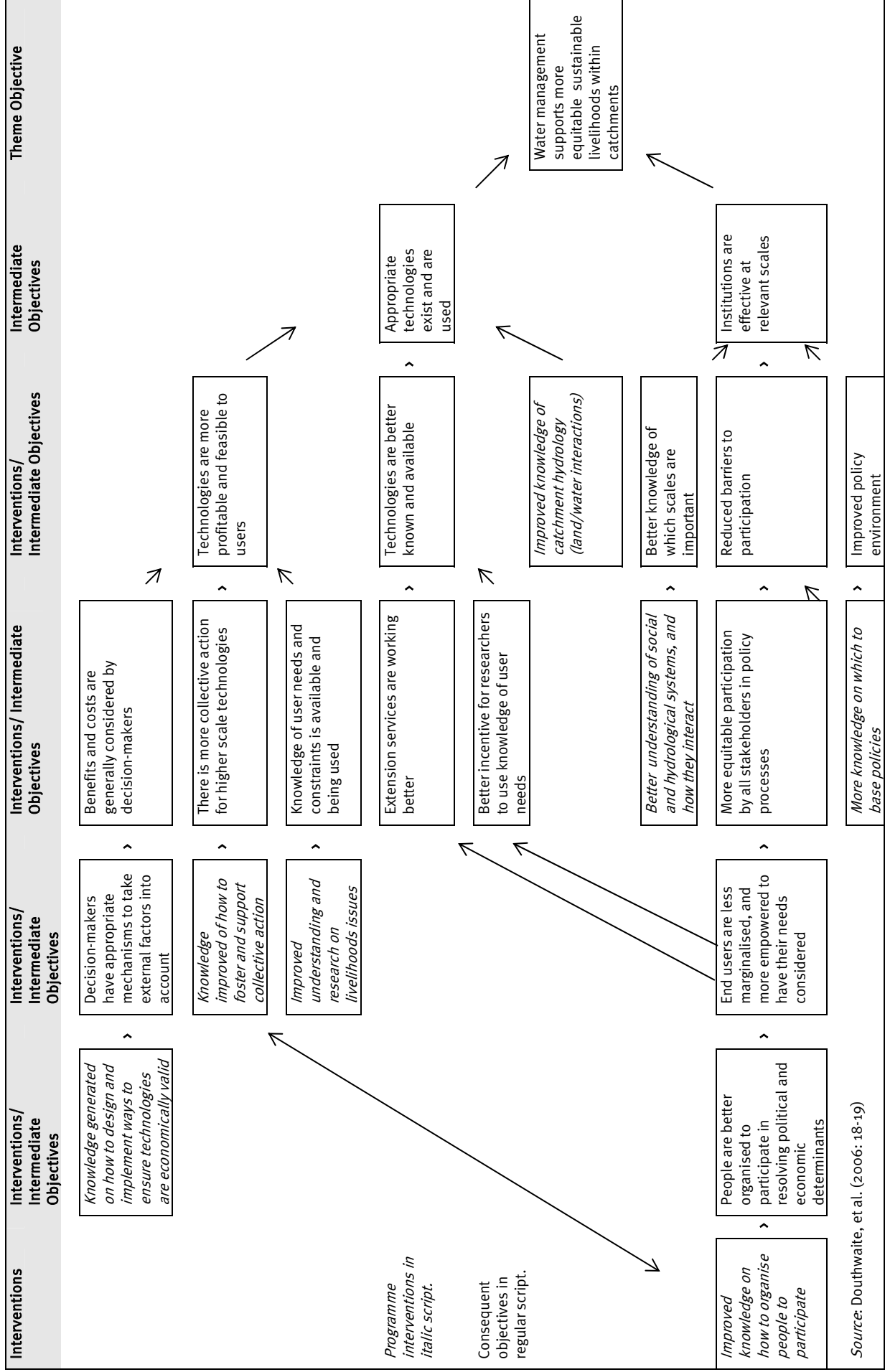
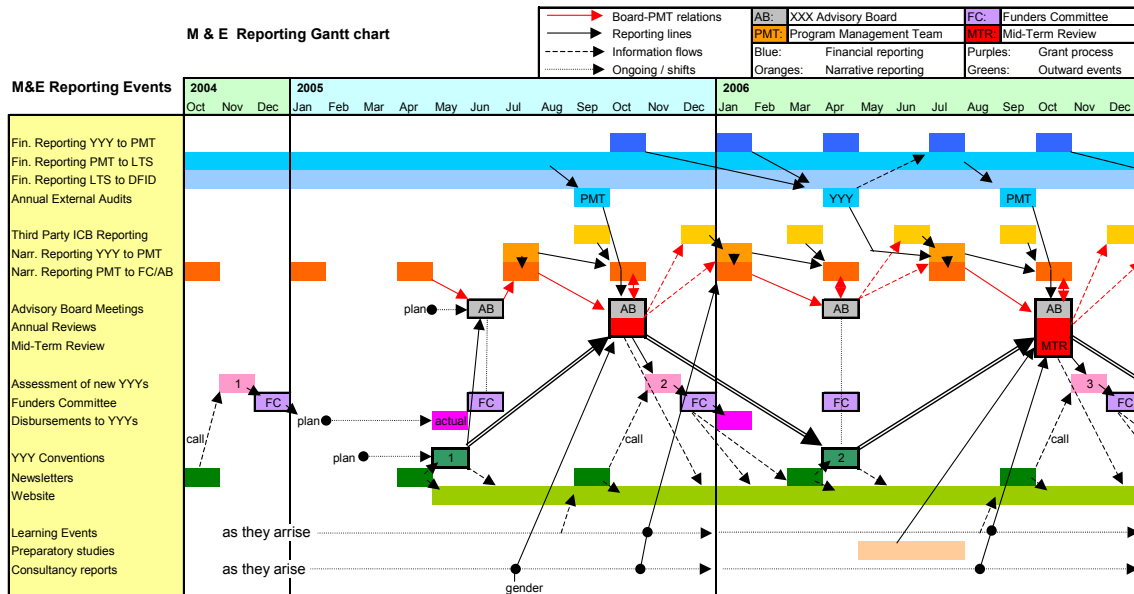


Figure 6 (continued). Example of Objective Tree from the CPWF: Water and People in Catchments



For the *Project Timeline*, important external events should also be included, such as key policy events or meetings. An example of a Gantt chart is given below.

Figure 7. Example of a Gantt chart



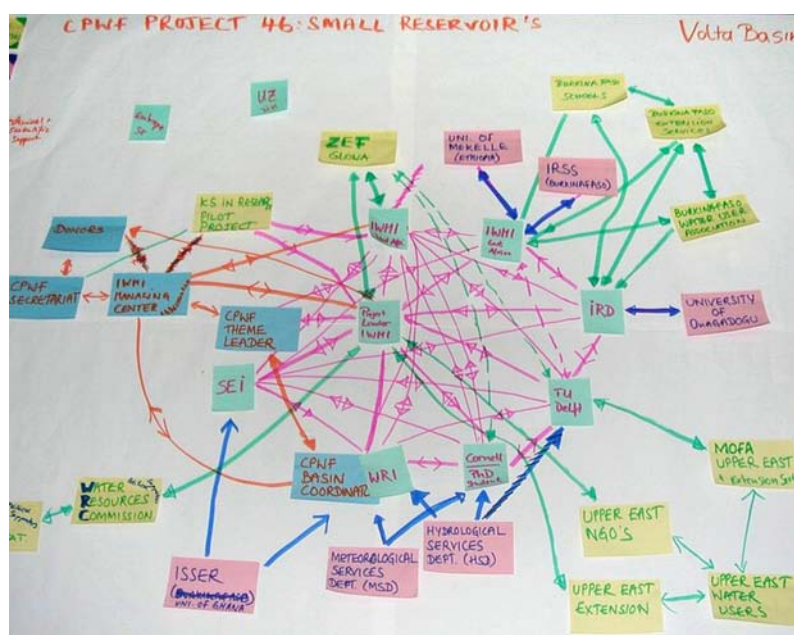
Source: Davies (2005: 27)

For the *Network Map*, see an example below in Figure 8 (and see also Figure 4 in the section above on SNA).

Douthwaite, et al. (2006: 7) suggest that it may be useful to draw both the programme’s social network as it exists now, and what the desired network will look like two years after the end of the programme. The difference between the current and future network then becomes a good evaluation and planning tool. The networks that have been drawn up can then be plotted into SNA software such as NetDraw (see section above on SNA).

The CG Science Council (the body responsible for the overall quality and relevance of science in the CGIAR) requires that all CPWF’s Medium Term Plans are reported in terms of impact pathways, and this reflects the fact that impact pathways is a tool that is especially well suited to self-assessment, providing a space for reflecting on the current situation, and for adjusting future plans, e.g. at a mid-term review.

Figure 8. In the process of constructing a Network Map for one of the CPFW projects



Source: Douthwaite (2006)

Box 1. CGIAR, including SPIA, ILAC, LTI, CIAT, IFPRI and ACIAR

The Consultative Group on International Agricultural Research (CGIAR) has done both institution-wide as well as project-specific work on M&E of policy research. They produced a sourcebook on performance measurement of research institutions and programmes in 2003 (CGIAR Secretariat, 2003) and a planning document the following year (CGIAR Working Group on Performance Measurement, 2004). The Sourcebook presented examples of institutional performance assessments from the public and private sector, including the Common Assessment Framework (CAF) used by the EU; Balanced Scorecards; Total Quality Management; and ISO 9000. It also presented a survey of qualitative (peer-review based) and quantitative approaches to performance measurement within the higher education sector. Overall they found that the main assessment methods of research performance used were bibliometric analysis and peer review – the relevant example from the UK is the Research Assessment Exercise (RAE). The CGIAR Working Group on Performance Measurement did not wish to adopt any of these models wholesale, but rather chose elements to design their own CGIAR-wide **Performance Measurement System**. The selected core elements of this evaluation system are presented in Appendix 1.

Moving down from the system-wide level to the institutional, programme and project levels within the CGIAR system, there are several innovative models that are being tested. At system level, the CG Science Council supports the Standing Panel on Impact Assessment ([SPIA/www.sciencecouncil.cgiar.org/activities/spia/index.html](http://www.sciencecouncil.cgiar.org/activities/spia/index.html)), CGIAR supports the Institutional Learning and Change Initiative ([ILAC/www.cgiar-ilac.org](http://www.cgiar-ilac.org)), and the International Centre for Tropical Agriculture (CIAT) supports the Learning to Innovate Group (LTI). These institutional platforms and communities of practice have in turn supported individual CGIAR institutions, programmes and projects in carrying out specific M&E approaches of their policy research.

Examples include **Appreciative Inquiry** (Acosta and Douthwaite, 2005), **Horizontal Evaluation** (Thiele, Devaux, Velasco and Manrique, 2006), **Innovation Histories** (Douthwaite, 2005) and **Episode Studies** (Leksmono, et al., 2006). (These four approaches have all been given sections in this paper below.) The International Centre for Tropical Agriculture (CIAT) has been active in exploring new M&E approaches, and has used both Innovation Histories (at project level) and Appreciative Inquiry (at an institutional level).

For the large-scale Challenge Programme on Food and Water (CPFW), an evaluation and learning model called **Impact Pathways** has been developed (described in the section above). This approach combines a logical model

(such as a logframe) with a social network model. Mid-term reviews of all CPWF projects are now carried out using the Impact Pathways approach and are reported in this form to the CG Science Council.

In addition, two of the CGIAR institutions should be singled out for attention, namely IFPRI and ACIAR. The International Food Policy Research Institute (IFPRI) in Washington DC has been working on impact assessment of its policy research since the mid-1990s. They have found that a **case study approach** is the most appropriate to articulate, measure and document the impacts of their policy research. Beginning in 1998, they undertook a series of case studies that examined relevant policy processes and the use of information by policymakers. They set up an *Impact Assessment Discussion Paper* series to publish their studies (www.ifpri.org/impact/impact.htm). These case studies are similar to the Innovation Histories mentioned above. They identified a policy research project and then traced the stages it went through and the impact it had on its surroundings. The institutional evaluation approach used by IFPRI is outlined further in Appendix 1.

The Australian Centre for International Agricultural Research (ACIAR) in Canberra has specialised in **economic impact assessment** of its policy research projects and programmes. Their framework for economic impact assessment is explained in Raitzer and Lindner (2005). Economic impact, as assessed for agricultural research, is often a quantification of the value of productivity improvement or inputs saved. It is often taken as granted that such productivity increases will foster gains throughout the broader target economies, and thereby achieve ultimate goals of poverty alleviation. Thus, the values presented are, in most cases, quantifying intermediate indicators, rather than benefits to the poor. Like IFPRI, ACIAR has hosted its own published *Impact Assessment Series* since 1998 (www.aciar.gov.au/web.nsf/doc/ACIA-5KL9S2), where all impact studies undertaken to date can be found.

2.1.4 Modular Matrices

Level: Projects, Programmes

Another self-assessment tool for policy research programmes that also draws on the perspective of Social Network Analysis is the ‘modular matrix’ approach developed by Rick Davies. This approach is designed to help describe the internal linkages of a project or programme (Davies, 2005). While the approach may be easier to carry out for a defined research *project*, it can also be used for policy research *programmes*. For example, Davies has used this approach in his work with the Civil Society Partnerships Programme (CSPP) at ODI, and it forms the basis of the CSPP M&E framework.

The approach focuses on exploring how the components of a project or programme relate to one another – e.g. how the project’s outputs relate to its desired impacts, how its outputs relate to its stakeholders, or how its outputs relate to key future events. The matrix approach that he proposes is primarily descriptive. It can therefore be a useful tool for a mid-term review, when a research project wishes to describe and assess its current status, and think about how to move forward.

Davies presents numerous examples of matrices in his paper (Davies, 2005). Three of them that may be particularly relevant to policy research projects are presented here: (1) An Outputs x Impacts matrix; (2) an Outputs x Stakeholders matrix; and (3) a Gantt chart.

(1) **An Outputs x Impacts matrix** gives the *desired* contribution of each project output to one or more of the project’s intended impacts. A fictional example is given below.

(2) **An Outputs x Stakeholder matrix** assesses to what degree each of the project’s outputs is reaching one or more of the project’s stakeholders or target audiences. A fictional example is given below.

Figure 9. Example of an Outputs x Impacts matrix

Impacts \ Outputs	Strengthen local research capacity on topic	Increase awareness about topic among policymakers and in media	Build relationships between research partners and civil society organisations	Influence change towards more pro-poor policy
Project launch		XXX		X
Website	X	X		X
One-on-one meetings with policymakers		XXX		XX
Public meeting series	X	X	XXX	X
Network building	XX	X	XXX	X
Research reports	XXX		X	
Policy briefs	XX	XXX	X	XX

Figure 10. Example of an Outputs x Stakeholder matrix

Stakeholders \ Outputs	Research partners	National policymakers	Bilateral and multilateral donors	Civil Society Organisations	Media
Project launch	X	XX	X	XX	XXX
Website	XX	X	XX	XX	XX
One-on-one meetings with policymakers	X	XXX			
Public meeting series	XX	X	X	XX	X
Network building	XXX		X	XXX	
Research reports	XX		X	X	
Policy briefs	X	XXX	XX	X	XX

The matrix shows Outputs vs Stakeholders (the groups of actors that the project wishes to engage with). For each output, crosses are distributed across the output row depending on where the output's *desired* impact lies. The matrix can then be compared to the project's *actual* distribution of effort across different groups of actors, in order to assess whether any resources need to be reallocated.

(3) A **Gantt chart** can be used in a variety of ways, including as a matrix that plots outputs against key events in the future. An example has already been given in the section on Impact Pathways above.

Davies considers it most useful to use a modular matrix approach when several matrices are worked out in parallel: 'These matrices can be seen as building blocks, which when combined with each other, can help create a larger construction of what is going on' (2005: 17). If a research project or programme developed the three matrices presented here, for example, then staff would have a description of some key internal linkages: to what extent their outputs (past, current or planned) contributed to their desired impacts; to what extent their outputs were geared towards their target audiences; and to what extent their outputs were aligned with significant events (e.g. policy events or key meetings). They could then assess whether they were pleased with the current status, or whether they would like to shift their internal efforts towards improved outputs (e.g. a better website) or different outputs (e.g. a meeting series in addition to the printed publications) in order to better achieve their desired impacts, reach their target audiences, or connect with key events in the future. In addition, the group work

required to put together the matrices may in itself be a useful tool to clarify project or programme processes and to provide a space for reflection on how well these processes are working.

2.2 Performance Area II: Evaluating management

This section gives examples of approaches that have been used to monitor and evaluate to what extent a policy research project/programme/institution is well managed, i.e. whether its internal systems and processes enable it to produce high-quality work that contributes to its intended strategy.

2.2.1 'Fit for Purpose' Review

Level: Projects, Programmes, Institutions

DFID commissioned an evaluation of the Renewable Natural Resources Research Strategy (RNRRS) programmes in 2003-04, carried out by the Performance Assessment Resource Centre (PARC). RNRRS is a cluster of ten parallel DFID-funded research programmes in the area of natural resources, operating from 1995-2005. Collectively, the ten programmes amounted to more than 1,600 research projects. PARC chose to focus on three areas for evaluation: programme governance and management; the quality of the science; and the impact on poverty (LTS, Noragric and OPM, 2005). Their evaluation of the area of programme governance and management is most relevant to this section.

In order to evaluate programme governance and management, the evaluators adopted the basic approach of reviewing whether the programme's processes were 'fit for purpose', i.e. whether the processes that were in place were appropriate to the programme's stated purposes, which were focused on poverty reduction, livelihoods strengthening, capacity building, skills transfer and influencing activities. In order to address the question of whether the programme's processes were 'fit for purpose', the evaluation team combined document review with a comprehensive list of interviews. They interviewed all DFID and programme managers. They reviewed planned and actual management roles, responsibilities and practices; they examined both the planning processes as well as the plans of the programmes, including log frames and other strategic documents; they looked at management structures, especially focusing on decision-making, communications, reporting, monitoring, and human and financial resources management. Through this approach of interviewing and reviewing, they were able to compare the programme's actual processes with its stated purposes.

Box 2. The UK Department for International Development (DFID)

The UK Department for International Development (DFID) has done various work in the area of M&E of policy research. Some of this is in the form of guidelines to the research projects that they fund – especially those funded by DFID's Central Research Department (CRD). In the M&E guide issued to all CRD-funded research projects, they stress the importance of using a **logframe** with SMART indicators (Specific, Measurable, Attainable, Relevant and Timing). All research projects are required to report to DFID on the basis of their project logframe in annual reports (in addition to six-monthly financial reports).

More innovative work has been done in DFID's **externally commissioned evaluations** of their own research programmes. Three examples are the evaluation of DFID's Renewable Natural Resources Research Strategy (LTS, Noragric and OPM 2005) described in the section immediately above; the evaluation of DFID's Engineering Knowledge and Research (EngKaR) Programme (Arnold, Court, Stroyan and Young, 2005); and the evaluation of DFID's Crop Post-Harvest Research Programme (CPHP) 'Partnerships for Innovation' (Barnett, 2006).

The EngKaR programme was evaluated by ODI and Technopolis. They were asked to evaluate the programme's effectiveness, quality, relevance, dissemination, uptake and impact. They did this through a mixture of desk-based reviews of project documents and data, a questionnaire survey of project leaders, interviews with DFID staff, independent project reviews, in-country visits, and a communications review. Through this **multi-pronged approach** they were able to bring to light both internal process gaps in the programme as well as some of the programme's successful relationships and synergies that might otherwise have been missed.

The evaluation of the CPHP ‘Partnerships for Innovation’ projects, carried out by The Policy Practice, drew on **National Systems of Innovation (NSI)** by assessing whether and how projects contributed to successful innovation. Projects were evaluated through visits and interviews, where information was gathered about: The relationships surrounding the project (their partners and stakeholders, quality of partnership); a range of possible project successes, including scientific papers, coalitions, uptake, changes, and impacts; the project’s uptake model; whether the project had changed as a result of a DFID shift towards partnership; and any lessons the project staff themselves shared. Through the underlying emphasis on NSI, the evaluation was able to draw out the importance of the relationships, systems and environment that the projects found themselves in, and the extent to which they had managed to contribute to this environment and to become an effective part of it, thus ensuring uptake of project findings.

2.2.2 ‘Lighter Touch’ Quality Audit: Expert visits

Level: Projects, Programmes, Institutions

The term ‘Quality Assurance’ has become linked to the M&E field over recent years. For information on recently set-up Quality Assurance processes in the World Bank, UNDP and Danida, see Haden (2005). In this section the Quality Assurance process of a UK university will be presented.

Universities in the UK are periodically evaluated on the basis of their research outputs, in the Research Assessment Exercise (RAE). However, they are also periodically assessed on the basis of their internal processes, to determine whether they are fit to continue as public research bodies, contributing knowledge and skills to students and to the public arena more broadly. This assessment is carried out by the Quality Assurance Agency for Higher Education (QAA). The QAA aims to assess a university’s internal management processes, decision-making processes, and quality assurance processes. Their key question is to determine whether these processes are of appropriate standards for a research body operating in the public sphere. The approach they take in order to answer this question is termed a ‘lighter touch’ institutional audit (QAA, 2006). The core components of the approach are as follows – and here the School of Oriental and African Studies (SOAS) in London will be used as an example:

SOAS has its own Quality Audit and Assurance Committee (QAAC), which coordinates the process. The school submits a briefing document to the QAA called an Institutional Briefing Paper, with supporting documentation. This allows the school to demonstrate which processes and systems it has in place to carry out its own (self-)monitoring. The student body also contribute a document. The QAA then arranges for a group of four academics to visit the school twice. Their first visit is a three-day ‘briefing visit’, where an overview is sought, and where the group considers all documentation submitted, and meets with senior managers and with a group of student representatives. At the end of that visit, they decide on two departments which they will ‘audit trail’ in their second visit. The second visit is the ‘main visit’ and takes place over a week. The evaluation group is briefed on how the school itself monitors and evaluates the two departments that they have chosen to look at more closely (at SOAS this self-monitoring is done through the internal Departmental and Programme Review process), and the group then assesses whether the institution is conducting and managing that process rigorously. They do this through more documentation review, relating specifically to the two departments, as well as a range of interviews with faculty, staff and students in the two departments. During their main visit they also examine any broader institutional matters that may be relevant.

At the end of the process they give the school a judgment of confidence (or limited confidence or no confidence), and whatever recommendations they believe would enhance the rigour of the school’s internal processes.

All in all, this evaluation approach gives more responsibility to the research institution itself, since the institution has to demonstrate that it is carrying out its own processes of monitoring and evaluation. At the same time, the expert visits will provide an opportunity for the institution to reflect and answer concrete questions about its own processes, and may also provide an opportunity to make a few changes and improvements.

Box 3. The European Centre for Development Policy Management (ECDPM)

The European Centre for Development Policy Management (ECDPM) in Maastricht recently commissioned an **external evaluation** of the institution as a whole (Matter, Mwai, Sefuke and Sherriff, 2006). The evaluation was undertaken by a four-person team, composed of two experts from Europe (including the team leader) and two from the ACP (both from Africa). For further details on the evaluation, see Appendix 1.

In its own work on evaluation, ECDPM is involved in **the EU's 3Cs initiative** (www.three-cs.net). The 3Cs initiative is a joint initiative by the Heads of Evaluation for External Cooperation of the EU Member States and the European Commission, with a view:

- To explore and assess the role played by the Maastricht Treaty precepts, *coordination*, *complementarity* and *coherence* (3Cs), in the EU's development co-operation policies and operations; and
- To determine how far these have been applied in practice and with what impact.

The organisations involved are ministries in Belgium, France, Germany, Ireland, the United Kingdom, Sweden, the Netherlands, the European Commission, the European Centre for Development Policy Management (ECDPM), and Contactivity (an organisation working on ICTs).

The initiative is in the process of carrying out a series of evaluation studies, which are expected to produce evidence, lessons and recommendations to strengthen the quality and effectiveness of European development assistance (for an example, see ECDPM and ICEI, 2005).

ECDPM also hosts **the Pelican Initiative**: the Platform for Evidence-Based Learning and Communications for Social Change. The Pelican Initiative is an active dgroup mailing list. The central question addressed by the list is: How can we learn more from what we do while at the same time having the biggest possible impact on the social change processes in which we engage? For more information, go to www.dgroups.org/groups/pelican/index.cfm?CookieTested=TRUE.

2.2.3 Horizontal Evaluation: Visits from colleagues for mutual learning

Level: Projects, Programmes, Institutions

Selected CGIAR projects have piloted an evaluation approach that is in some respects similar to the institutional quality audit used for UK universities. CGIAR have called it 'horizontal evaluation' (Thiele, Devaux, Velasco and Manrique, 2006). Horizontal evaluation is similar to a 'lighter touch' institutional audit in that the programme under review is expected to carry out some form of self-assessment, and present this to the evaluators, who come to visit. The difference is that in the horizontal evaluation approach, the evaluators are not regarded as experts, but rather as colleagues from other programmes who are working on similar themes, and who have an interest in learning from the programme evaluation. Thus the evaluation visits turn into meetings where mutual learning takes place. Both the programme staff and the visitors make recommendations for how to strengthen the programme.

Thiele, Devaux, Velasco and Manrique (2006) sum up their experience with this approach as follows:

'Our experience indicates that the following factors are critical for the success of a horizontal evaluation:

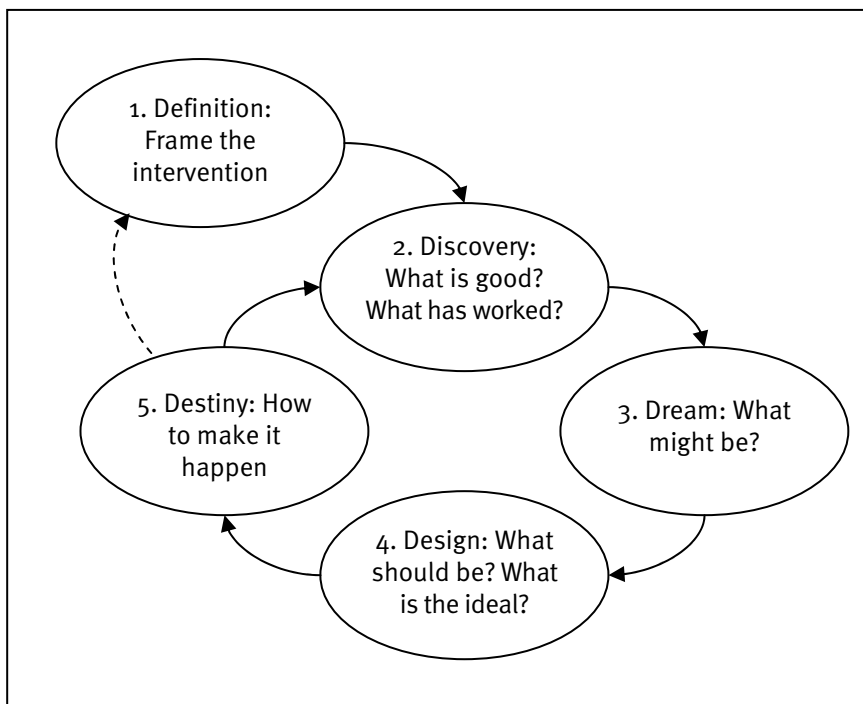
- Selecting a moment for the [evaluation] workshop when the new R&D methodology is sufficiently advanced so that there is real substance to review but not so finished that there is little space for modification.
- Careful selection of visitors to ensure that they have diverse perspectives, possess adequate knowledge and experience, and are perceived as 'peers' rather than 'superiors.'
- Good facilitation to create an environment of trust, to focus the attention of participants on the methodology under review ... and to manage time adequately.
- A limited number of clearly defined evaluation criteria.
- Well prepared presentations and field visits to ensure that visitors have the information to understand adequately the methodology under development.'

2.2.4 Appreciative Inquiry

Level: Projects (though only long-term projects), Programmes, Institutions (though only smaller institutions, or parts of institutions)

Appreciative Inquiry is a method of self-evaluation and learning based on the assumption that we have just as much to learn from our successes as from our failures. The method was first developed in the 1980s, but has since been tried out in the CGIAR-affiliated International Centre for Tropical Agriculture (CIAT) (Acosta and Douthwaite, 2005). Appreciative Inquiry is typically carried out in a three-day workshop with the institution's staff (though it is also possible to make it shorter), and it follows five steps.

Figure 11. The Appreciative Inquiry '5-D' model



Source: Acosta and Douthwaite (2005: 2)

1. **Definition:** Be clear about the specific aspects of the institution that will form the focus of the inquiry. For the CIAT Appreciative Inquiry workshop, the following three themes were chosen: (i) Fostering innovation; (ii) Cooperative relations; and (iii) Effective communication and knowledge sharing.
2. **Discovery:** Staff interview each other one-on-one, asking the interviewee to tell a story about a time when he or she felt that the institution or programme was at its best. At the CIAT Appreciative Inquiry workshop, the following questions were used (Acosta and Douthwaite, 2005: 3):
 - *Fostering innovation:* Reflect on your time at CIAT or other experiences. What were the highpoints for you working to foster innovation? Select one highpoint, at a time when you felt happiest and most alive, when you felt you were doing creative, useful, meaningful work that really made a difference. What were you doing, what felt good, who else was involved and what did you feel you achieved?
 - *Cooperative relations:* Identify and describe a scenario that demonstrates the positive aspects of working together, cooperating to get something done. Who and what was involved, why did it work, what were you doing and what were other people doing?
 - *Effective communication and knowledge sharing:* What different types of communication occur between CIAT's different geographic locations? What do you value most about effective

communication? When does this happen for you? Who and what is involved in the best types of communication? Why is effective communication good for you and for work on fostering innovation?

The focus on highpoints is meant to direct thinking away from problems and deficits, and towards opportunities and assets. Once people start to reflect more closely on their highpoint stories, they will also start to form thoughts about the combination of resources, opportunities and relationships that contributed to it. The interviews should try to capture stories rather than opinions, since the events and relationships of the stories will form the basic material for the learning.

After the one-on-one interviews, participants come together in small groups and share key findings, themes, and ‘gems’ (key quotes) from their interviews. Common themes are then presented back to the group.

3. **Dream:** Staff are asked to imagine a scenario 5 or 10 years in the future, when the ‘highpoint’ experiences of the institution or programme have become the norm. How would they describe and explain this to e.g. a journalist? Participants share their answers in small groups and common themes are then presented back to the group.
4. **Design:** Staff are then asked to think about what would need to happen for the institution or programme to reach the ideal future scenario. Working in small groups, they formulate their answers in the form of ‘provocative propositions’. Provocative propositions are bold, affirmative statements, in the present tense, that are grounded in the group’s collective history, yet stretch, challenge and interrupt the status quo. The propositions should relate to the themes that have been chosen. They are then shared with the group.
5. **Destiny:** This is the final phase, where evaluation and learning are turned into concrete plans and actions. The nature of the plans, and who is chosen to follow up on them, will vary depending on the specific institution or programme. If the approach has worked well, and the staff buy into the dreams and provocative propositions that they have formulated, then they will already be motivated to build change into their own agendas.

The Appreciative Inquiry approach is most valuable for smaller institutions, or as a mid-term review tool at programme level, and it can be followed up by further Appreciative Inquiry workshops at regular intervals. It can be helpful in fostering a learning culture. As Acosta and Douthwaite (2005: 3) point out: ‘Ultimately, one will know if AI has really taken hold when the ‘appreciative eye’ is used with increasing frequency, beyond the initial pilot phase.’

2.3 Performance Area III: Evaluating outputs

This section gives examples of approaches that have been used to monitor and evaluate the research outputs produced by a policy research project/programme/institution. Outputs are understood to be the tangible goods and services that a research project/programme/institution produces (e.g. Working Papers, journal articles, policy briefs, website, meetings, events, networks, etc).

2.3.1 Evaluating academic articles and research reports (quality of science criteria)

Level: Projects, Programmes, Institutions

The evaluation of the DFID-funded RNRRS research programmes (LTS, Noragric and OPM, 2005), as mentioned above (in the section on ‘Fit for Purpose’ Reviews), examined three areas. The second of these areas was ‘quality of science’. In order to assess the quality of the science research carried out

and produced by the 1,600 research projects under the 10 RNRRS programmes, the evaluators drew heavily on traditional academic means of assessment. In their report they say:

Quality of science includes aspects related to the correct formulation of hypotheses, the appropriateness of scientific inputs, research methodologies and processes, and research outputs and outcomes. The most objective, and most commonly used measure of quality of science is the quality of publications in refereed journals. (LTS, Noragric and OPM, 2005: 10)

This is an example that shows that while policy research programmes may not always wish to adopt too heavy academic frameworks or modes of assessment, they may choose to incorporate certain academic criteria in order to evaluate certain outputs – such as journal articles or research reports. The criteria that the RNRRS evaluation chose to incorporate are given below. A group of specialists was then asked to provide individual assessments on aggregated research outputs, based on these criteria, giving a score of 1 (none/worst) to 10 (much/best) for each.

'Quality of science' criteria (LTS, Noragric and OPM, 2005: 11):

- To what extent the programme contributes to new knowledge;
- To what extent the programme uses existing knowledge creatively in new contexts;
- Rating of the programme in relation to its innovation and scientific risk-taking with comment on projects that are innovative and projects that are not;
- Demonstrate awareness of all current knowledge (journals, books, web-based information) including in developing country literature, English language literature and non-English language literature;
- Extent to which the expected science achievements outlined in the log frame have been met (key projects, outputs at programme level);
- Extent to which projects and the programme have contributed to science capacity building in the scientific communities in developing countries;
- Development of long-term institutional relationships between UK institutions and Southern institutions;
- Rating of the overall result knowledge dissemination from programme:
 - To science community (refereed, non-refereed, web-based, other media);
 - To developing country policy audiences;
 - To developing country outreach services;
 - To developing country end users (farmers, foresters, fisher folk);
 - To the international donor community.

2.3.2 *Evaluating policy and briefing papers*

Level: Projects, Programmes, Institutions

Policy and briefing papers should be assessed against different criteria than academic journal articles. Policy papers are written specifically for the purpose of using evidence to shed light on a policy area. Briefing papers are produced with the same purpose, but may be much shorter (perhaps 1-6 pages). Young and Quinn (2002) argue that good-quality policy and briefing papers have three core components: (i) they say what the *problem* is; (ii) what the possible *solutions* are, including the author's preferred solution; and (iii) what policy *recommendations* follow from this.

When evaluating policy and briefing papers, these three components can be assessed in turn:

(i) Assess whether the policy or briefing paper has made it clear from the beginning what the (policy) problem is. There should be both a claim and evidence to support it. There should also be a coherent argument reflected, for example, in the following features (Young and Quinn, 2002):

- Clear, descriptive section titles and numbering. Do not label simply label the introduction 'Introduction'; instead, give it a descriptive title.
- The opening sentence of each section drives the argument.
- The first (or last) sentence of each paragraph makes the most important point.
- Effective use of paragraphing (for those readers who only skim).
- Coherence within the text.

These features have one key aim in common: they make it easy for someone to skim the paper and still absorb its overall argument.

(ii) Assess whether the policy or briefing paper has made it clear what the possible solutions are. Policy options should be succinctly outlined and compared. The preferred policy option should be highlighted.

In the policy options element, the policy adviser needs to show his or her expertise and take the lead in the argument to strongly advocate for his or her chosen option. Remembering that policy science should be problem-oriented and targeted, this is the opportunity for you to prove that yours is a practical solution to the outlined problem, and therefore a valuable contribution to the policy debate and the policy community in general. (Young and Quinn, 2002)

(iii) Assess the final conclusion and recommendations section. Many readers will only read this section, along with the introduction or executive summary. It is therefore important that the section will be able to give readers a clear overview of the whole paper. Questions for this section include (Young and Quinn, 2002):

- Does the section synthesise only the major findings of the study?
- Are the recommendations logically divided into separate measures and clearly presented?
- Are all recommendations effectively written?
- Does the conclusion provide a sense of completeness to the paper?

2.3.3 *Evaluating websites*

Level: Projects, Programmes, Institutions

The International NGO Training and Research Centre (INTRAC) in Oxford has developed guidelines on how one might evaluate websites that aim to communicate research to an international audience (Taylor 2001). This is the output of a DFID-funded project in which INTRAC evaluated the usage of eight websites that aim to disseminate policy research on urban development issues. (The eight websites were: Community Development Society <http://comm-dev.org>, European Network for Sustainable Urban and Regional Development Research www.european-association.org/ensure, Forum: Habitat in Developing Countries www.forumhabitat.polito.it, Global Development Research Centre www.gdrc.org, Network Association of European Researchers on Urbanisation in the South www.naerus.org, Resource for Urban Design Information www2.rudi.net/rudi.html, United Nations International Environmental Technology Centre www.unep.or.jp, DFID-Urbanisation www.lboro.ac.uk/garnet/UrbanKaR/DFID-KAR-URBAN.html.)

The evaluation team decided to make conclusions regarding key qualities associated with constituent components of each website. They defined the constituent components as follows (Taylor 2001:2):

- **Architecture** refers to the structure of the website and the logic by which the pages interconnect. This component is examined in order to assess the site's *navigability*.
- **Technology** refers to issues around the quality of the code and the appropriateness of any technologies used. This component is assessed in relation to how *accessible* and *available* the site will be to an international audience with different software and hardware capabilities.
- **Style** refers to the *appearance* of the website and issues around the layout and display of text and images.
- **Content** refers to the quality, authority, readability, relevance and timeliness of text and images, and the degree to which *user interaction* is supported.
- **Strategy** refers to the degree to which the site has met *stated objectives* concerning its target audience or market.
- **Management** relates to the human and financial resources that the site has at its disposal.

These components are described in more detail in the paper. The evaluation team then used a combination of the following methods in order to address their questions:

- *Observation*: Some evaluation questions could be answered by the evaluator whilst browsing the website concerned, paying special attention to e.g. directory structure, downloadable documents and 'meta-tags' (keywords that help promote a website in search engines).
- *Automated Tests*: The evaluators ran a series of automated tests, including HTML (Hypertext Mark-up Language) validator, available free from <http://validator.w3.org/>; Test for the compatibility of technology with particular browsers, using the website analysis tool 'Dr HTML'; Analysis of a single webpage was run for free at www2.imagiware.com/; Check the status of each website's links, using Xenu's Link Sleuth at <http://completelyfreesoftware.com/>; Test for compatibility errors between each website's code and particular browsers, using the browser compatibility table available at <http://hotwired.lycos.com/webmonkey/browserkit>; Test to measure the download time for each home page – using a 28,000 baud modem – using the free website analysis tool *Bobby*, www.cast.org.
- *Usability Tests*: Testing of users undertaking specified tasks is considered to be one of the best methods for evaluating websites. For the eight websites evaluated, five INTRAC colleagues volunteered to complete a task-based questionnaire. Their objective was to record how many clicks it took them to locate a specified piece of information. This figure was then compared to the minimum number of clicks necessary to find that information.
- *A User Survey*: An online survey was carried out in order to gather information about the profile and experiences of users located internationally.
- *A Webmaster Survey*: The webmasters of the sites were asked to complete a short questionnaire for their website concerning log-file data (e.g. number of hits, page requests, unique visitors, etc), and a number of internal organisational issues.
- *Telephone Interviews*: A semi-structured interview was carried out with a member of each website team in order to develop an understanding of the organisational processes by which website content was selected and edited.

2.3.4 Evaluating networks

Level: Projects

The International Institute for Sustainable Development (IISD) has developed guidelines for how to evaluate knowledge networks (Crech, 2001). IISD defines 'knowledge networks' as networks that aim to share information and create new knowledge, strengthen research and communication capacity among network members, and identify and implement strategies to engage decision-makers more

directly, linking to appropriate processes in the areas of policy and practice. These are the types of networks that policy research projects or programmes may wish to participate in or establish. How can such networks be evaluated?

Using components from several planning and M&E tools, including SWOT (strengths-weaknesses-opportunities-threats) analysis, logical framework analysis, and Outcome Mapping, IISD have created three stages for network evaluation:

(1) **Planning:** used at the beginning of network activities, to record the work plan, the beneficiaries of the work (partners and stakeholders), and the indicators of change desired for major projects or programs of work within the network, and for the network as a whole. Creech (2001: 16-22) provides a detailed table that can be worked through during the planning phase. In sum, some of its key questions are:

- What can members contribute to, as well as receive from, the network?
- What will success look like for the network as a whole?
- For each activity, who is going to benefit, be changed or influenced by the work?
- What will be the indicators of success for each activity?

(2) **Monitoring: Progress journals.** Creech (2001) suggests that the use of progress journals is the best way to keep track of ongoing network activities and developments. Each network member keeps a journal where main activities are recorded, as well as any feedback received or any interesting or unusual events. The activities recorded do not need to be assessed. The journal can be shared e.g. quarterly with the network coordinator.

(3) **Evaluation:** An annual evaluation gives the network a chance to assess whether the network's component programs are on track, whether anticipated outcomes are being achieved, and whether adjustments need to be made to activities, objectives, work plans, and expected outcomes. The evaluation builds on the progress journals, paying special attention to the activities that were successful and those that were not, and to any interesting stories or unexpected opportunities that have come up. Creech (2001: 25-27) gives a detailed example.

The annual evaluation should then be used to assess whether the network as a whole is realizing its potential. Creech (2001) suggests examining the following questions:

- Is the network linking effectively into relevant policy processes; is the level of recognition and influence of the network and its members increasing within these circles?
- Are members adding value to each others' work, and creating new work together that might not have happened otherwise?
- Is there an exchange and building of capacity across the network membership?

2.3.5 *After Action Reviews*

Level: Projects

The After Action Review (AAR) is a simple tool to facilitate assessment of a task or activity that has been carried out (Ramalingam, 2006). In a policy research project it could be used to evaluate e.g. a meeting or an event. It works by bringing together a team to discuss the event or activity in an open and honest fashion. The systematic application of properly conducted AARs across a programme or institution can help drive organisational change. As well as turning unconscious learning into tacit knowledge, it helps to build trust among team members and to overcome fear of mistakes. When applied correctly, AARs can become a key aspect of the internal system of learning and motivation.

There are many different ways to conduct AARs. The simplicity at the heart of the tool means there is much potential to experiment with the process and find the right ways that will work best with the group and the work item under review. The whole process should be kept as simple and as easy to remember as possible. The essence of the AAR is, however, to bring together the relevant group to think about a project, activity, event or task, and pose the following simple questions.

Figure 12. After Action Review questions

Question	Purpose
What was supposed to happen? What actually happened? Why were there differences?	These questions establish a common understanding of the work item under review. The facilitator should encourage and promote discussion around these questions. In particular, divergences from the plan should be explored.
What worked? What didn't? Why?	These questions generate reflection about the successes and failures during the course of the project, activity, event or task. The question 'Why?' generates understanding of the root causes of these successes and failures.
What would you do differently next time?	This question is intended to help identify specific actionable recommendations. The facilitator asks the team members for crisp and clear, achievable and future-oriented recommendations.

A Retrospect follows the AAR format, but involves asking the following more detailed questions:

- What did you set out to achieve?
- What was your plan to achieve this?
- How did this change as you progressed?
- What went well and why?
- What could have gone better?
- What advice would you give yourself if you were to go back to where you were at the start of the project?
- What were the two or three key lessons you would share with others?
- What next for you in terms of this project?
- Can you think of a story that summarises your experience of work on this project?
- What should we have learned from this project a year from now?
- Are there any lessons for you personally?

Key points and practical tips (Ramalingam, 2006):

- Post the questions up on flipchart sheets prior to the session, with answers then written on the sheet as the session progresses. The completed sheets can then be stuck up around the room to serve as a reminder of the progress.
- Participants are participants, not a passive audience. The facilitator should prepare leading questions and may have to ask it of several people. The questions can be asked on an individual or a team basis. The team mechanism is ideal, but if suggestions are slow coming, the facilitator could go around the room asking each individual to express one thing that worked and one thing that did not.
- If there are issues with either openness or time, it may be worthwhile to gather ideas first and then facilitate the discussion in the group environment.
- Ideally, an uninvolved note-taker should be asked to minute the session. This will enable better capture of the learning.
- The actionable recommendations should be as specific as possible. For example, an AAR

following a workshop could have the following recommendation: ‘Make more time to understand the audience.’ A better SAR would be ‘Make contact with the organising body representative and ask about the range of participants before planning the workshop.’

- Participants of an AAR should include all members of the team. A facilitator should be appointed to help create an open environment, promote discussion and draw out lessons learned.
- AARs should be carried out immediately, while the team is still available and memories are fresh. It is recommended that AARs be incorporated at key points during a project, activity, event or task in the early planning stage, although they are often completed at the end.
- AARs can be conducted almost anywhere, and will vary in length. For example, a 15-minute AAR can be conducted after a one-day workshop, or a much longer meeting could be held to reflect on the strategy development process throughout a large organisation.

2.4 Performance Area IV: Evaluating uptake

This section gives examples of approaches that have been used to monitor and evaluate the uptake of a policy research project/programme/institution, i.e. the extent to which its research and recommendations have been ‘picked up’ by others. Uptake is here defined as direct responses to the research project/programme/institution (e.g. the research is mentioned in a government policy paper, on a range of websites, referred to in a newspaper article, etc).

2.4.1 *Impact Logs*

Level: Projects, Programmes

In the Research and Policy in Development Group (RAPID) at ODI, impact logs are used to keep track of some of the direct responses that the research outputs trigger, and this in turn informs programme evaluation. An impact log is a list of the informal feedback, comments, and anecdotes that a programme receives from people who have encountered or used its research outputs. It is not a systematic way of assessing user perceptions; rather, it is a way of capturing the qualitative and non-systematic feedback on research outputs that would otherwise get lost. As the Impact Log grows longer, the cumulative effect can be valuable in assessing where and how the project or programme is triggering the most direct responses, and in informing future project/programme choices.

2.4.2 *New Areas for Citation Analysis*

Level: Projects, Programmes, Institutions

A more systematic way of seeking out direct responses to research outputs is to use an expanded form of citation measurement. As mentioned in the introduction, policy research programmes will not usually use conventional academic citations in peer-reviewed journals as a primary monitoring and evaluation tool. They may, however, wish to combine academic citation analysis (if they have any traditional academic outputs) with new citations measurements, such as those developed by Grant Lewison at City University, London. Drawing on Lewison (2005), the following list of six types of documents/texts can be used to analyse and trace the impact of policy research outputs:

- *International standards*, such as those published by the EU. Most of these can be found on the internet.
- *Government policy documents*.
- *Operational guidelines* issues by government bodies or professional associations. NGO guidelines would fall into this category.

- *Training manuals* and, in some cases, textbooks.
- *Newspaper articles*.
- *Websites*. Use a search engine to pick up references to policy research outputs on other websites, including mention of the research, link to the research programme website, or link to downloadable document.

When carrying out citation analysis, several steps can be taken, depending on how comprehensive the analysis needs to be:

- The first step is to search the documents and texts listed above, gather and count citations to one's own research programme and its outputs, and assemble these into a table.
- A second step might be to also gather citations to key competitors or other major players in the field, in order to compare and assess one's own level of citations with them.
- A third step might be to combine the citation counting with some more qualitative assessment. For example, Lewison (2005: 1528) suggests that mention of policy research in newspaper articles can be further analysed and evaluated using the following criteria: the amount of attention given to the research in the article; the size and prominence of the news article; the tone used (optimistic, neutral, critical); and the newspaper's circulation and the socio-economic profile of its readership.

2.4.3 User Surveys

Level: Projects, Programmes, Institutions

A more proactive way to gather information about responses to research outputs is to carry out a user survey. As this is a long-standing M&E tool, it will not be presented in detail here. However, it is relevant to note that user surveys range from large-scale questionnaire-based data gathering exercises to small focus groups. The use of focus groups can in some cases be viewed as an appropriate alternative to (academic) peer review in policy research programmes, since it may be more accurate to glean the perceptions and reactions of a range of users rather than 'experts' when it comes to evaluating, for example, a website or a policy brief.

2.5 Performance Area V: Evaluating outcomes and impacts

This section gives examples of approaches that have been used to monitor and evaluate the outcomes and impact of a policy research project/programme/institution. Outcomes and impacts are here defined as changes in behaviour, knowledge, policies, capacities and/or practices that the research has contributed to, directly or indirectly (e.g. a change in government policy implementation, a change in working practices among NGO practitioners, a reduction of poverty in a certain area, strengthened livelihoods, strengthened civil society input into policy processes, etc).

2.5.1 Outcome Mapping

Level: Projects, Programmes

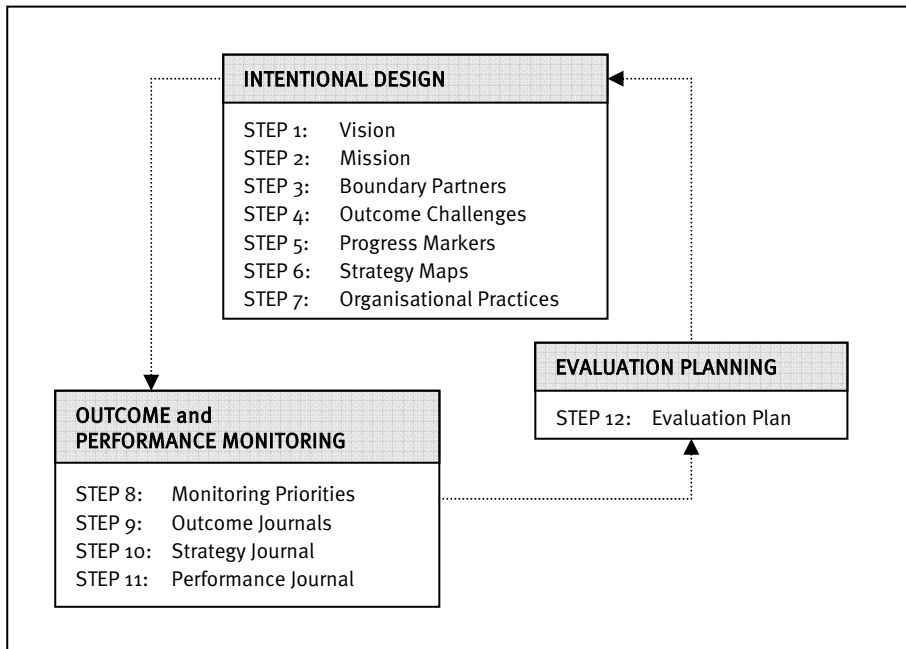
Outcome mapping is an M&E tool developed by the International Development Research Centre (IDRC) (Earl, Carden and Smytulo, 2001). It introduces M&E considerations at the planning stage of a project or programme, and, if followed through, it can serve as a valuable evaluation mechanism. The approach moves away from the notion that M&E are 'done to' a project or programme, and, instead, actively engages the team in the design of a monitoring framework and evaluation plan and promotes self-assessment throughout.

Its recent success is based on its explicit emphasis on *relationships* and on *change*. It focuses on the following key words:

- **Behavioural change:** The ‘outcomes’ of a research programme are defined as changes in the behaviour, relationships, activities, or actions of the people, groups, and organisations with whom the programme works directly. These outcomes can be logically linked to a programme’s activities, although they are not necessarily directly caused by them.
- **Boundary partners:** Those individuals, groups, and organisations with whom the programme interacts directly and with whom the programme anticipates opportunities for influence. Most activities will involve multiple outcomes because they have multiple boundary partners.
- **Contributions:** By using Outcome Mapping, a programme is not claiming the achievement of development impacts; rather, the focus is on its contributions to outcomes. These outcomes, in turn, enhance the possibility of development impacts – but the relationship is not necessarily a direct one of cause and effect.

The originality of the methodology is its shift away from assessing the changes in state (for example, policy relevance, poverty alleviation, or reduced conflict) and toward changes in the behaviours, relationships, actions or activities of the people, groups, and organisations with whom a policy research programme works directly. In other words, Outcome Mapping establishes a vision of the human, social, and environmental betterment to which the programme hopes to contribute and then focuses M&E on factors and actors within that programme’s direct sphere of influence. Outcome Mapping does not belittle the importance of changes in state (such as cleaner water or a stronger economy) but instead argues that for each change in state there are correlating changes in behaviour. Outcome Mapping is done in three stages, as illustrated below.

Figure 13. The three stages of Outcome Mapping



Source: Earl, Carden and Smytulo (2001)

(1) **The first stage, Intentional Design,** helps a programme establish consensus on the macro level changes it will help to bring about and plan the strategies it will use. It helps answer four questions: Why? (What is the vision to which the programme wants to contribute?); Who? (Who are the program’s boundary partners?); What? (What are the changes that are being sought?); and How? (How will the programme contribute to the change process?).

(2) **The second stage, Outcome and Performance Monitoring**, provides a framework for the ongoing monitoring of the programme's actions and the boundary partners' progress toward the achievement of outcomes. It is based largely on systematised self-assessment. It provides the following data collection tools for elements identified in the Intentional Design stage: an 'Outcome Journal' (progress markers); a 'Strategy Journal' (strategy maps); and a 'Performance Journal' (organisational practices).

(3) **The third stage, Evaluation Planning**, helps the programme identify evaluation priorities and develop an evaluation plan.

The process for identifying the macro-level changes and designing the monitoring framework and evaluation plan is intended to be participatory and, wherever feasible, can involve the full range of stakeholders, including boundary partners. This can in itself be an important tool for relationship building, and can encourage future ownership and use of findings.

Box 4. The International Development Research Centre (IDRC)

The International Development Research Centre (IDRC) in Canada has been one of the pioneers in the field of evaluating policy research, mainly through their development of the **Outcome Mapping** approach (Earl, Carden and Smytulo, 2001)(described in the section above). Outcome Mapping focuses on relationships between actors, and how changes occur in these relationships and in the behaviour of the actors involved.

When it comes to its own institutional assessment, IDRC uses a **decentralised evaluation system** (IDRC Evaluation Unit, 2005; 2006) with a utilisation focus (Patton, 2002). For more details on this evaluation system, see Appendix 1.

2.5.2 RAPID Outcome Assessment

Level: Projects

RAPID Outcome Assessment (ROA) was developed as part of the Process and Partnership for Pro-poor Policy Change (PPPPC) project carried out by the CGIAR-affiliated International Livestock Research Institute (ILRI) and the Research and Policy in Development Group (RAPID) at ODI (Leksmono, et al., 2006). The PPPPC project seeks to identify and institutionalise innovative research and development mechanisms and approaches that lead to pro-poor policy. The project is a response to the need to better understand the processes and mechanisms that lead to pro-poor decisions at the policy level, and aims to provide recommendations to improve the impact of work by ILRI and its partners.

ROA draws on elements from Outcome Mapping, and was designed as a learning methodology to assess the contribution of a project's actions and research on a particular change in policy or the policy environment. It is a flexible, visual tool that can be used to map changes in the project and its environment, and it can be used in conjunction with other evaluation tools and methods to evaluate a particular project or programme.

The key steps of the ROA were originally designed as follows:

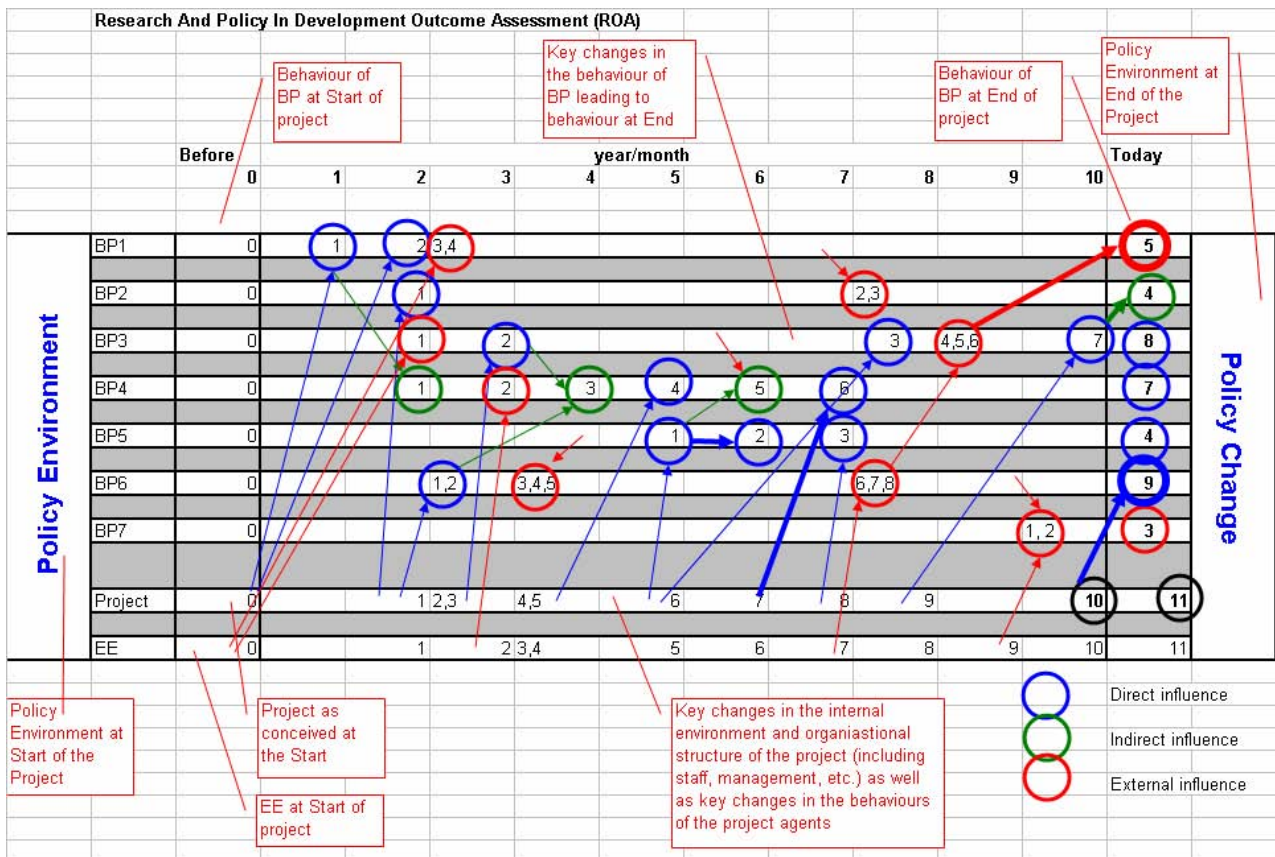
1. Describe the policy environment at the end of the project;
2. Describe the policy environment at the beginning of the project;
3. Identify the key policy actors or agents of change;
4. Within the agents of change, identify the boundary partners that are conducive to the change or that influence the policy environment;
5. Describe the behaviours of the boundary partners that are conducive to a change in the policy environment or policy;
6. Describe the behaviours of the boundary partners at the beginning of the project;
7. Map the key changes in behaviour for each boundary partner from the start of the project;

8. Map the key changes in the internal environment of the project including organisational changes, outputs and changes in behaviour during the same period;
9. Map the external influences including the actions of strategic partners and other exogenous factors during the same period;
10. Determine the level of impact/influence of the project on the changes in behaviour of the boundary partners;
11. Determine the level of impact/influence of external influences on the changes in behaviour of the boundary partners and the project;
12. Refine the conclusions with in-depth interviews and assess the real contribution of the project to the policy environment;
13. Write report.

The intention was that steps 1–11 could be covered in a workshop with key stakeholders to produce a table similar to the one below.

However, after some discussions with researchers, concerns were raised about the ability of some workshop participants less exposed to training methodologies to understand the ROA methodology and to follow the discussion. It was decided that the initial workshop with project related staff and advocacy CSOs would not follow the steps described above rigidly. Instead, various participatory training techniques were used to collect the information needed to complete the expected output of ROA.

Figure 14. Example of ROA Output



Source: Leksmono, et al. (2006)

In the ROA workshop, the following steps were actually carried out:

1. *Description of the policy environment at the beginning and the end of the project.* Participants were asked to write on cards what they think has changed in the smallholder dairy sector in Kenya in the last few years. The cards were then organised into following categories: organisational change, policy change, practical change and behavioural change.
2. *An introduction to the Smallholder Dairy Project (SDP) which would be used as the case study.* A presentation about the history, organisation and activities of the SDP was given by the original project manager from the relevant Ministry.
3. Identification of SDP project time line, policy context and external factors. Participants were divided into groups:
 - a. Group 1 (SDP Staff) developed a project time-line and identified the key events and changes;
 - b. Group 2 (Other Participants) were asked to develop a 'rich description' of the policy context for smallholder dairy production in Kenya currently and in 1996; and to identify any key external events which influenced how the policy context for smallholder dairy production has changed.
4. *Identification of key players.* Participants were asked to name all the key players in the development of the smallholder dairy sector in Kenya. They were then asked to select the three that they thought had had the most impact on the smallholder dairy sector, and to write the reasons why on a coloured card – yellow for their first choice, green for their second choice and blue for the third choice. A score was calculated for each stakeholder where a yellow card equalled three points, a green card two points and blue card one point.
5. *Description of key actor behaviour.* Participants were divided into groups to consider specific groups of stakeholders and wrote on cards their behaviour now, their behaviour in 1996, key points when their 'behaviour and attitudes' changed in between. On the back of those cards, they wrote why they thought that behaviour change happened, and what impact they thought that the change has had on others. The cards were then pinned to the wall.
6. *Participants were then asked to look at all the factors on the wall and see if they could see any links between them.* Each link identified was connected using a piece of string and the reason was noted. A copy of the cards and links was projected via Power Point onto another wall. There was not enough time to complete this activity on the second day; a smaller group completed the map the following day.
7. *Identification of issues* to be followed-up in the interviews with key players.

In the end, the workshop was able to produce a table similar to the example above and important information about the project was gathered.

From the process of applying ROA in the workshops, some useful findings and experiences should be noted:

- It is important to know in advance who will be participating in the workshop: their background, their expectations, their relationship with the project, and their agendas, as this will help smooth the facilitation process and enable better planning of workshop activities;
- The participants should remain the same throughout the workshop – people who are not available to attend for the duration of the workshop should not be invited as it will delay the progress of the workshop and disrupt the small group discussions;
- It is important that all participants understand the concepts of policy, behaviour and attitude, in order to be able to accurately identify changes in policies. The facilitator should ensure that the participants understand before moving on to the next activity;
- Involving the project staff in the planning of the workshop, and also involving them in the facilitation process, is useful for focusing the activities, and also enables one to gain 'insights' into participants' statements and comments, thus enabling a better understanding of the project;

- Planning needs to be done to avoid a long plenary discussion, especially when identifying links between key events. The background of each participant should be noted so they can be asked to assemble in different groups of key actors and asked to discuss the links in small groups. It would be useful for one group to discuss the link between the project and the external environment, including the key events, and the other group to discuss links between other key actors and the key events. Another idea is to give three sets of strings to each participant: blue for direct effects from the project; green for effects from the key actors; and red for effects from the external environment. Participants are then asked to use the strings to link two key events and note why they think there should be a link. The results should then be discussed in a plenary session;
- Writing a daily report and having it available for the next day of the workshop is very useful in focusing the discussion on the second day.

2.5.3 *Most Significant Change*

Level: Projects, Programmes, Institutions

The Most Significant Change (MSC) approach involves the collection of significant change (SC) stories, and the systematic selection of the most significant of these stories by panels of designated stakeholders or staff. By recording, collecting, reviewing and choosing between SC stories, staff at all levels gain greater awareness of the kinds of impacts that the project, programme or institution is working towards. This focused attention encourages a form of ongoing and indirect monitoring of the work carried out. MSC also gives a project, programme or institution a better understanding of whether and how it is achieving its purposes. In addition it provides the project, programme or institution with a set of valuable PR materials.

MSC is a popular technique, and many adaptations have already been made. These are discussed in Davies and Dart (2005). A generic description of the process can be outlined as follows:

1. *How to start and raise interest:* The first step in MSC generally involves introducing a range of stakeholders to MSC and fostering interest and commitment to participate.
2. *Defining the domains of change:* The next step is to identify the domains of change to be monitored. This involves selected stakeholders identifying broad domains – for example, ‘changes in people’s lives’ – that are not precisely defined like performance indicators, but are deliberately left loose, to be defined by the actual users.
3. *Defining the reporting period:* The third step is to decide how frequently to monitor changes taking place in these domains.
4. *Collecting SC stories:* SC stories are collected from staff across the programme/institution. The stories are collected by asking a simple question such as: ‘During the last month, in your opinion, what was the most significant change that took place as a result of your work?’ It is initially up to respondents to allocate their stories to a domain category. In addition to this, respondents are encouraged to report why they consider a particular change to be the most significant one.
5. *Selecting the most significant of the stories:* The stories are then analysed and filtered up through the levels of authority typically found within a programme or institution. Each level of the hierarchy reviews a series of stories sent to them by the level below and selects the single most significant account of change within each of the domains. Each group then sends the selected stories up to the next level of the hierarchy, and the number of stories is whittled down through a systematic and transparent process.
6. *Feeding back the results of the selection process:* Every time stories are selected, the criteria used to select them are recorded and fed back to all interested stakeholders, so that each subsequent round of story collection and selection is informed by feedback from previous

rounds. The organisation is effectively recording and adjusting the direction of its attention – and the criteria it uses for valuing the events it sees there.

7. *Collecting the stories in a document:* After this process has been used for some time, such as a year, a document is produced with all stories selected at the uppermost organisational level over that period in each domain of change. The stories are accompanied by the reasons the stories were selected. In the case of projects or programmes, the programme funders are asked to assess the stories in this document and select those that best represent the sort of outcomes they wish to fund. They are also asked to document the reasons for their choice. This information is fed back to managers.
8. *Verification of stories:* If desired, the selected stories can then be verified by visiting the sites where the described events took place. The purpose of this is two-fold: to check that stories have been reported accurately and honestly, and to provide an opportunity to gather more detailed information about events seen as especially significant. If conducted some time after the event, a visit also offers a chance to see what has happened since the event was first documented.
9. *Quantification:* If desired, quantification can take place at two stages. When an account of change is first described, it is possible to include quantitative information as well as qualitative information. It is also possible to quantify the extent to which the most significant changes identified in one location have taken place in other locations within a specific period.
10. *Secondary analysis and meta-monitoring:* The next step is monitoring the monitoring system itself, which can include looking at who participated and how they affected the contents, and analysing how often different types of changes are reported.
11. *Revising the system:* The final step is to revise the design of the MSC process to take into account what has been learned as a direct result of using it and from analysing its use.

One example of a collection of SC stories is *Perceptions and Practice: An anthology of impact assessment experiences* (Sayce and Norrish, 2006). The authors have not explicitly used the MSC approach, but they have gathered stories of significant changes among the projects of the Technical Centre for Agricultural and Rural Cooperation ACP-EU (CTA), and have written them up as narratives. In this form, the stories serve both as points of learning as well as a chance for CTA to share its experience with others.

2.5.4 Innovation Histories

Level: Projects, Programmes

A similar approach to MSC is the recording of Innovation Histories. This method has been developed by the CGIAR-affiliated International Centre for Tropical Agriculture (CIAT) (Douthwaite and Ashby, 2005). In their field of policy research, one of the primary goals is to enable rural innovation, and yet it is difficult to monitor and evaluate to what degree their research projects and programmes achieve this goal. They have therefore found it valuable to record and analyse stories of innovation when they do occur, and to use these as learning and evaluation tools.

- The first step in the recording of an innovation history is for people who have been involved in the innovation to jointly construct a timeline of the innovation history, based on their recollections and on available documents. It is important to note that in order to record an innovation history, there must be a clear innovation or change to focus on. The process of preparing this history stimulates discussion, reflection and learning amongst stakeholders.
- The participants then construct two or more actor matrices for selected points in the timeline to capture the dynamics of changing relationships. An example of an Actor x Actor matrix is given below.

Figure 15. Example of an Actor x Actor matrix

	Actor A	Actor B	Actor C
Actor A		Relation of A–B	Relation of A–C
Actor B	Relation of B–A		Relation of B–C
Actor c	Relation of C–A	Relation of C–B	

1. Identify and list actors for a phase of the innovation history.
2. Actors may be NGOs, donors, etc.
3. Draw matrix describing type of relationship (collaboration, funding, etc).
4. Identify relationships that were: a) crucial; b) problematic; or c) absent but needed.

Source: Douthwaite and Ashby (2005: 2)

The matrices can be converted into network maps using a social network mapping programme such as InFlow or Pajek for easier visualisation (see section above on Social Network Analysis).

- Participants then decide on key aspects or themes of the innovation history, and this will determine which further literature they need to gather, and which stakeholders they need to interview. The interviewees are asked about additional information regarding events, any events or people who should be added to the history, their own perception of what happened, and which themes or relationships they see as most important.
- The material gathered is then brought together into a written account (sometimes referred to as a ‘learning history’). This can contribute to project or programme evaluation, learning, and to plans for future projects/programmes. Douthwaite and Ashby (2005: 4) recommend the following format for these accounts:
 - *Introduction*: describes the background to the innovation and the rationale for creating the innovation history and explains why this approach is useful.
 - *Methodology*: describes the framework used and the data-gathering methods.
 - *Case study or studies*: this is the meat of the report; the narrative describing what actually happened based on the timeline and actor network maps.
 - *Discussion and conclusions*: describes the factors that fostered and constrained the innovation process. These findings are compared with existing literature – in particular that relating to the view that innovation is an interactive and experiential learning process mitigated by social networks.
 - *Synthesis*: compares and contrasts the main findings from each case study (if there is more than one), or discusses the implications of the findings for the project or future similar projects/programmes.
- Finally, on the basis of the initial detailed account of the innovation history, more concise informational products can be prepared that summarize the innovation process for wider dissemination of findings. These may include public awareness materials, policy briefs or articles in professional journals.

2.5.5 Episode Studies

Level: Projects, Programmes

The Research and Policy in Development Group (RAPID) at ODI has over the past few years been using an approach called Episode Studies to map the direct and indirect contributions of policy research to policy changes. This approach takes into account the fact that an enormous range of different factors influence most policy processes, waxing and waning and in different combinations over time. Looking

for the impact of only one factor (such as research) is therefore quite difficult. Episode Studies address this difficulty by ‘tracking back’ from policy changes to identify a range of key actors and decisions and assessing the relative importance of different factors, including but not limited to policy research. The key difference between Innovation Histories and Episode Studies is that Innovation Histories ‘track forwards’ from an innovation to its impact, while Episode Studies ‘track back’ from a policy impact or a policy change to look at the range of factors that may have influenced it. The approach is carried out as follows:

- *Constructing a narrative:* An Episode Study starts by constructing an historical timeline leading up to the observed policy change in question. This involves interviewing a range of stakeholders who have been involved in the process, creating a timeline of key policy decisions and practices, along with important documents and events, and identifying key actors.
- *Assessing the relative role of research:* The next step is to explore why those policy decisions and practices took place and to assess the role of research in that process. This can be done through further interviews with key actors, reviewing the literature and crosschecking conflicting narratives. Research questions used in the RAPID Episode Studies included:
 - To what extent was the impact of research on policy-making shaped by political and institutional structures and ideological assumptions?
 - To what extent did local involvement, the quality of research and communications strategies affect the impact that research had on policy-making in particular areas?
 - To what extent did researchers and policy-makers share particular kinds of networks, common goals and chains of legitimacy for particular policy areas?
- *Analysis:* Key factors that facilitated or hindered research uptake can then be analysed. In the RAPID Episode Studies these factors were grouped into the four broad areas of the RAPID Framework: The political context; The use of evidence; Links and relationships; and External factors (Crewe and Young, 2002).
- *Lessons:* Finally, lessons can be drawn from the study about how researchers can strengthen their ability to inform and influence policy processes and changes.

For examples of Episode Studies published by RAPID, see Young, Kajume and Wanyama (2003), Buchanan-Smith (2003), Christiansen with Hovland (2003) and Solesbury (2003). Episode Studies have also been used in the Process and Partnership for Pro-poor Policy Change (PPPPC Project) that RAPID is working on together with the CGIAR-affiliated International Livestock Research Institute (ILRI) (Leksmono, et al., 2006).

Box 5. Overseas Development Institute (ODI)

The Research and Policy in Development Group (RAPID) at the Overseas Development Institute (ODI) has been working on research-policy linkages since 2002. RAPID’s main contribution to the debate has been the development of **the RAPID framework** (Crewe and Young, 2002), which looks at how research uptake may be enabled or disabled by various factors in the areas of Context (political institutions, political system, etc), Evidence (quality, availability, communication, etc), and Links (relationships, networks, trust, etc).

The RAPID framework has been applied to numerous qualitative **Episode Studies**, where the various factors leading up to a policy change have been mapped and analysed (for a list of case studies, see the description of Episode Studies above). RAPID has also merged the use of rapid Episode Studies with elements from Outcome Mapping, in order to capture the multiple actors and relationships involved in a policy change process. The result is the method of **RAPID Outcome Assessment (ROA)**, which has been described above.

ODI is also in the process of drawing up a comprehensive institutional M&E strategy. More information on this can be found in Appendix 1.

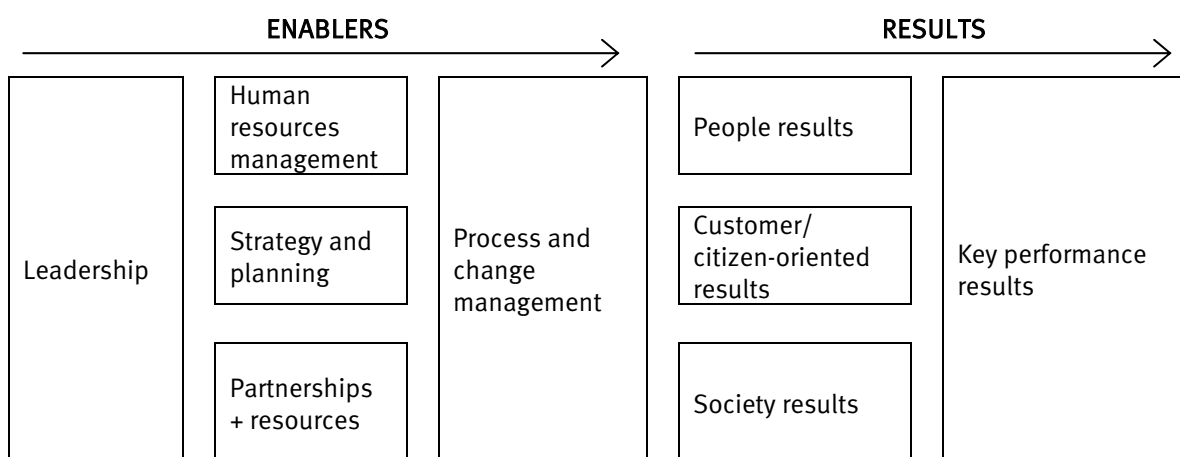
3 Institutional evaluations: Additional concerns

When conducting *institutional* evaluations of policy research institutes or think tanks, as opposed to project or programme evaluations, there are clearly a few additional concerns that need to be addressed. Most importantly, these include:

- how to assess the organisation's governance structure and accountability;
- its leadership (including overall management structure, management capacity, and quality of management);
- its administration (including administrative systems, capacity, IT systems, use of physical resources and space);
- its human resources (including recruitment procedures, appraisals, staff development, etc);
- its finance systems; and
- its environment (including the political economy, the history, the cultural spheres, and the field of relationships that the institution is operating within).

This survey paper has only focused on M&E of policy research, rather than M&E of organisational governance, leadership, administration, human resources, finances or environment. However, for those particularly interested in these areas, a good place to start to look is the CGIAR Secretariat's (2003) *Performance Measurement of Research Institutions and Research Programs – A Sourcebook*, which presents four conceptual frameworks for assessing overall institutional leadership: the Common Assessment Framework (EU); Balanced Scorecards; Total Quality Management; and ISO 9000. For more information on the **Common Assessment Framework (CAF)**, which is a Total Quality Management (TQM) tool used by the EU, see the leaflet prepared for the 2002 Second Quality Conference (Danish Ministry of Finance 2002). An outline of the CAF model is given in Figure 16 below. Each of the areas comes with a set of questions.

Figure 16. The CAF model



Source: CGIAR Secretariat (2003: 6)

Another good place to start is the website *Reflect and Learn* (www.reflectlearn.org), which has collected a number of organisational self-assessment models. The site has a separate section on organisational self-assessment in research centres, including organisational assessment tools from IDRC and the Innovation Network. The site also links to an organisational assessment model specifically designed for agricultural research centres, made available by the International Service for National Agricultural Research (ISNAR) (Peterson, Gijsbers and Wilks, 2003).

For policy research institutes and think tanks in the South, see James McGann's (2006) *Best Practices for Funding and Evaluating Think Tanks and Policy Research*, which gives examples of evaluations of think tanks located in the South, and recommends best practices for institutional evaluations. UNDP and its partners have also developed an evaluation process tool designed to fit the circumstances of CSOs (including research institutes) in the South, namely the **Participatory Organisational Evaluation Tool (POET)**. POET measures and strengthens seven capacity areas, shown in Figure 17 below.

Figure 17. The POET model

Capacity Area	Focus
Human Resource Management	staff development, recruitment, compensation (salary and benefits), personnel evaluation, and grievance and conflict resolution
Financial Resource Management	budgeting, forecasting, fundraising, and cash management
Equitable Participation	field-based programme practices related to project access and project benefit
Sustainability of Programme Benefits	the impact of environmental, economic, political, institutional, and cultural factors
Partnering	collaboration with other CSOS, donors policy makers, and private sector entities
Organisational Learning	teamwork, information-sharing and capacity for generating information that leads to improvement of current practice
Strategic Management/ Governance	board practices; planning practices; and, commitment to goals, mission and philosophy

Source: UNDP (1998: 6)

POET offers a way of discussing, measuring and strengthening each of the seven areas in Figure 17. For more information on POET, including the relevant questions and process, and ways of scoring each of the areas, see UNDP (1998), or <http://reflectlearn.org/POET.php>.

4 Conclusion: Designing an M&E plan for a policy research project, programme or institution

4.1 Adding value

As Davies, Nutley and Walter state (2005: 16), one size does not fit all when it comes to evaluation of policy research. The route chosen by a particular research project, programme or institution will vary depending on the stated aims of the research, on the time and resources available for evaluation, and on the institutional and political setting. It will probably be necessary to combine elements and approaches for different performance areas into an M&E plan, and some projects or programmes may need to take a more experimental approach. The aim should be to design an M&E plan that fits the overall direction and intentions of the research programme, and that adds value to it.

Watts (2005) suggests that when putting together an M&E plan, its value is enhanced when the focus is just as much on learning as on accountability. This is also the rationale behind IDRC's user-focused evaluation system (IDRC Evaluation Unit, 2006). It goes without saying that if staff feel that the M&E activities are helping them in their work, rather than judging them, they are more likely to participate in them and to use the lessons that come out of them.

Another way to make the M&E plan add value is to bring the donors on board (Hyatt and Kaplan, 2006). As Hyatt and Kaplan point out, donors do not generally have a comfortable relationship with failure. The process of drawing up an M&E plan, or of carrying out M&E activities, might therefore be used as a platform to build up trust with the donor, to speak about calculated risks that the project, programme or institution plans to take, and to explain which 'failures' the organisation is willing to risk. Similarly, an M&E process may present a chance to involve other key stakeholders and to strengthen relationships with them.

Finally, the M&E process can also result in valuable public relations materials. To take just one example, organisations that use the Most Significant Change (MSC) approach may take the opportunity to publish and disseminate the Significant Change stories once they have been collected (e.g. Sayce and Norrish, 2006).

Below are some rough guidelines on steps to take when drawing up an M&E approach for a policy research project, programme, or institution. Needless to say these steps and suggestions will need to be adapted to the specific situation and aims of the project or organisation in question. They build on the five performance areas (and the associated M&E approaches) that have been highlighted in this paper:

- *Performance Area I – Evaluating strategy and direction:* Logframes; Social Network Analysis; Impact Pathways; Modular Matrices
- *Performance Area II – Evaluating management:* 'Fit for Purpose' Reviews; 'Lighter Touch' Quality Audits; Horizontal Evaluation; Appreciative Inquiry
- *Performance Area III – Evaluating outputs:* Evaluating academic articles and research reports; Evaluating policy and briefing papers; Evaluating websites; Evaluating networks; After Action Reviews
- *Performance Area IV – Evaluating uptake:* Impact Logs; New Areas for Citation Analysis; User Surveys
- *Performance Area V – Evaluating outcomes and impacts:* Outcome Mapping; RAPID Outcome Assessment; Most Significant Change; Innovation Histories; Episode Studies

4.2 Best practice checklist for policy research projects

The appropriate M&E approach for a policy research project will depend on the project's scale, timeframe, budget, aims, and any conditions set by donors. Almost all the M&E approaches outlined in this paper can be used at project level. However, the key is to focus on one or two manageable and appropriate approaches, and then to complement these with elements of other approaches if needed.

A few examples of possible combinations are:

- Design a Logframe at the start of the project, and monitor the Logframe indicators twice: at a mid-term review and an end review. Complement this with collecting MSC stories twice during the second half of the project, so that proper emphasis is put on stories of *change*.
- Design a Logframe and Social Network Map at the start of the project, and then revise and strengthen these by using the Impact Pathways model at a mid-term review, so that the importance of finding and using potential *impact pathways* is emphasised.
- Use Outcome Mapping from the start of the project, and complement this with a mid-term Horizontal Evaluation visit from two or more of your boundary partners, so that the importance of the project's *relationships* is emphasised.
- Design a Social Network Map at the start of the project, together with an Output x Stakeholder matrix, and a Gantt chart. Complement this with regular After Action Reviews after each project event, and Appreciative Inquiry workshops every 6 months throughout the course of the project, so that proper emphasis is put on the importance of ongoing *learning and reflection*.

4.3 Best practice checklist for policy research programmes

Designing an M&E approach for policy research programmes necessarily requires a little more time and thought than at project level. However, the advantage of this is that the M&E approach can form a valuable component of working out and embedding the overall programme strategy across all projects.

Step 1: Lay the foundations

Ensure that the foundations for programme M&E are laid at the beginning. This can be done for example through:

- Holding a two or three-day workshop to design a Logframe and a Social Network Map at the start of the programme, so that both the logical and social strategy are given due thought.
- Holding a two or three-day workshop to design a collection of Modular Matrices that adequately capture programme intent (e.g. Outputs x Stakeholders; Outputs x Outcomes; Gantt chart).
- Holding a four or five-day workshop to go through the Intentional Design phase of an Outcome Mapping process (Vision; Mission; Identify Boundary Partners; Identify Outcome Challenges; Identify Progress Markers; Strategy Maps; Organisational Practices).

Step 2: Set up quality assurance mechanisms across all projects

Ensure that some quality assurance mechanisms are in place across all projects from the beginning. Relevant examples may include:

- Appropriate peer review or user-review processes are carried out for all publications.
- After Action Reviews or similar mechanism is used after selected types of events.
- An Impact Log is kept either at programme level or by every project.
- Independent website evaluation is commissioned every 3 years (or similar).

Step 3: Coordinate project and programme M&E

Decide on how to combine project and programme M&E. This can be done, for example, through:

- A flexible relationship, where projects are asked to carry out their own M&E in the most appropriate way. Programme M&E can then draw on and collate information from project M&E, as well as carrying out a separate process of programme M&E.

- A close relationship, where all projects are asked to integrate their M&E systems. A good example of this practice is CGIAR's Challenge Programme on Water and Food (CPWF), where all 50 projects are required to use the Impact Pathways model for their mid-term review. Another example is the MSC approach, where all projects would be asked to collect MSC stories at regular intervals. The results of the project M&E will then constitute the bulk of the programme M&E.

Step 4: Choose an appropriate mixture of self-assessment and external evaluation

Decide on an appropriate combination of self-assessment and external evaluation. Also decide on how the self-assessment will feed into the external evaluation (e.g. through a self-assessment document that is presented to the external evaluators), and how the external evaluation in turn will feed back into the programme's strategy (e.g. through a review workshop, or a revised Logframe, or a revised Social Network Map, etc).

Self-assessment may consist of:

- An aggregation of M&E results provided by projects every 12 months (when using the Impact Pathways model and/or the MSC approach).
- A synthesis of information recorded in the Outcome Journals, Strategy Journals and Performance Journals every 12 months (when using the Outcome Mapping approach).
- A mid-term review of the programme's Logframe indicators and the programme's Social Network Map (when using a combination of programme Logframe and SNA); or a mid-term review of the programme's Modular Matrices (when using a Modular Matrix approach).
- Appreciative Inquiry workshops held for all programme staff at regular intervals (e.g. every six or 12 months).

External evaluation will usually consist of some form of visit by external evaluators. The programme should decide on what kind of visit will be most appropriate (e.g. in-depth or 'lighter touch'; structured or open-ended; by evaluation experts or peers in other organisations). Some examples of external evaluation visits include 'Fit for Purpose' Reviews, Quality Audits, and Horizontal Evaluation. If possible, the external evaluation should be commissioned for a mid-term review rather than an end-of-programme review, since at a mid-term review it can feed directly into programme strategy.

Step 5: End-of-programme evaluation

End-of-programme evaluations will depend on a number of factors, such as the M&E approach chosen, the requirements of donors, and the programme's M&E budget. The most important point to bear in mind is that the end-of-programme evaluation should ideally feed into future programmes, and some creativity may be needed to ensure that this happens. Possible approaches may include holding an end-of-programme learning workshop where staff from similar programmes are invited; publishing and distributing a collection of MSC stories from the programme's work; publishing and distributing a few Episode Studies of programme impact; or at the very least making sure that the end-of-programme evaluation report is short, reader-friendly, and available on the web.

4.4 Best practice checklist for policy research institutions

The steps outlined in this section might be used to draft e.g. a three or five-year institutional M&E strategy. The suggested steps draw on the pool of current experiences outlined in Section 2, and on the examples of institutional evaluations given in Appendix 1. They build on the five Performance Areas identified in this paper.

Step 1: Choose a decentralised or centralised focus

Decide whether the primary focus of the institutional M&E is decentralised or centralised. If it is decentralised, it will primarily focus on collecting strategic evaluations from all the institution's programmes and collating these (as done by e.g. IDRC – see Appendix 1). If it is centralised, it will

primarily focus on mapping out the institution's position in the field and strategic direction forward (as done by e.g. ECDPM – see Appendix 1). These focuses overlap a little, and can be combined (as done by e.g. CHSRF – see Appendix 1).

- If decentralised, decide whether to give each programme free reign over how they carry out and present their self-assessment, with only a minimum of common criteria (as in IDRC), or whether to ask each programme to present a self-assessment using a standardised format (as in CHSRF).
- If centralised, decide on how the programmes will be represented in the institutional M&E process, and whether they will contribute any programme self-assessment at all (and if so, in what form), or whether they will primarily contribute through helping to assess the overall direction of the institution.

Step 2: Decide on a mixture of self-assessment and external evaluation

Decide on an appropriate mixture of self-assessment and external evaluation. It is usually desirable to have an element of external evaluation at the institutional level, e.g. at the end of the M&E phase, which could be after three or five years. However, external evaluations will have no impact on an organisation unless a certain amount of self-assessment is also carried out.

Self-assessment:

- If a decentralised process for self-assessment has been chosen, then the primary locus for evaluation is at programme level (see the checklist for programmes above). Decide whether or not programmes should present their self-assessments in the same format, and how often (e.g. every 12 months).
- If a centralised process for self-assessment has been chosen, then generic approaches such as an institutional Logframe and an institutional Social Network Map may be easiest to work with. These may be drawn up for the entire M&E period (e.g. three or five years) and then reviewed every 12 months. It may be advantageous to use these in some combination, as this incorporates both a logical element (assessing whether the institution's outputs and activities are related to its intended outcomes) and a network element (assessing whether the institution is in a good position to interact and engage with key stakeholder groups in its field).

In addition, the institution should consider how to present and assess the management of 'non-research' units such as Finance and Human Resources.

Self-assessment should usually lead to a document (or documents) that can be presented to all staff, e.g. once a year, as well as to external evaluators in order to alert them to key areas for review.

External evaluation:

The institution should decide what kind of external evaluation it would gain the most from, and when it should take place (usually at the end of the M&E phase, though it may also be used as a mid-way review). Usually the external evaluation will entail a visit by a panel of evaluators. The institution will present its self-assessment document(s) to the evaluators, and they will then spend some time in the institution. They may review documents, interview staff, conduct focus group discussions, carry out participant observation, etc. Different types of visits include a 'Fit for Purpose' Review, a 'Lighter Touch' Quality Audit, a Horizontal Evaluation, or a self-designed variation. Good examples of external evaluation questions are given in the external evaluation report on ECDPM (see note on ECDPM in Appendix 1) and CGD (see Appendix 1). In part, the decision on an external evaluation will be determined by the institution's M&E budget and by any conditions set by its funders.

The institution may also decide to incorporate the tenets of Appreciative Inquiry into the evaluation visit (i.e. asking the external evaluators to explore the processes that work well, and how to build on these), or, in addition, the institution may decide to carry out a process of Appreciative Inquiry itself. This works best in small institutions.

Step 3: Note whether quality and uptake of outputs are monitored regularly

For an institutional evaluation it should not usually be necessary to carry out a review of the quality and uptake of individual institutional outputs. Rather, the important point during an institutional evaluation will be to assess whether adequate M&E processes are in place during the normal course of events to ensure that all outputs are monitored on a regular basis. An institutional M&E plan may note some or all of the following:

- Whether appropriate peer review processes are (or should be) carried out to monitor the (academic) quality of all research papers;
- Whether appropriate peer review and/or user-review processes are (or should be) carried out to monitor the (academic and user-friendly) quality of all policy briefs;
- Whether mechanisms are in place to capture citations of the institution's research in newspapers, on the internet, in government papers, and in academic works;
- Whether ad hoc comments on outputs are captured in Impact Logs or similar mechanism;
- Whether After Action Reviews (AARs) or similar mechanisms are carried out to monitor events or meeting series;
- Whether networks have their own M&E plan and carry it out;
- Whether a User Survey has been carried out recently to monitor the reach of publications, events, or the website, and/or whether another User Survey is necessary during the next M&E phase;
- Whether a separate evaluation of the website has recently been carried out, and its recommendations have been implemented, and/or whether another evaluation is desirable during the next M&E phase.

If the institution has decided to carry out processes of self-assessment, then it makes sense for the points above to be reviewed in the self-assessment document(s) produced e.g. every 12 months.

Step 4: Capture impacts

- If a decentralised M&E approach has been chosen, the institution should decide whether to ask each programme for specific information regarding impacts, and how often. Each programme may gather this information through e.g. Outcome Mapping workshops, RAPID Outcome Assessment workshops, Most Significant Change (MSC) stories, Innovation Histories, or Episode Studies. It will usually strengthen the evaluation process if this kind of information is gathered, and programmes may wish to go through this process e.g. every 2 years, or simply at the end of the M&E phase.
- If a centralised approach has been chosen, the institution will need to think creatively about how to capture its performance in this important area, and how often to do so.
 - A possible low-intensity approach may use a combination of the following: (a) Monitoring planned impacts against indicators in the Logframe at the end of the M&E phase; (b) Using the institution's Social Network Map to review the number of relationships that have been strengthened over the past M&E phase, and hence the current scope of the institution's possible impact; and (c) Extracting examples of actual impact from the institution's Impact Logs.
 - A possible high-intensity approach may involve collecting MSC stories from across the organisation at the end of the M&E phase, and going through a systematic selection process in order to arrive at the stories that most appropriately capture the kinds of impacts that the institution is aiming for.

Step 5: Use the evaluation to draw up a revised strategy for the next phase

At the end of the M&E phase (e.g. after three or five years), the institution will end up with some of the following documents:

- A collection of decentralised annual programme self-assessments (for an example, see IDRC Evaluation Unit, 2006);

- Or a collection of centralised annual institutional self-assessment document (for example, headings may include a combination of the following: Introduction and process; Review of institutional Logframe; Review of institutional Social Network Analysis; Review of management of ‘non-research’ units; Review of mechanisms for evaluating outputs and uptake; Examples of outcomes and impacts);
- A collection of MSC stories, or other stories of impact, which can be used as a publication in its own right;
- An external evaluation visit report (for an example, see the external report on ECDPM: Matter, Mwai, Sefuke and Sherriff, 2006).

These documents can then inform a new organisational strategy for the next phase.

4.5 When you are asked to evaluate someone else’s project, programme or institution

The best practice checklists given above are mainly meant to help staff who wish to build a more thoughtful and coherent approach to evaluating their own policy research project, programme or institution. However, we are also often asked to act as external evaluators of other policy research projects, programmes or institutions. In these cases, the evaluation usually has to be quicker and less comprehensive. Ideally, it should form one step of the project or institution’s own M&E plan. If it does not, however, it may be helpful (both to the external evaluators as well as to staff) to take the five performance areas outlined in this paper as starting points. This may be done, for example, by using a combination such as the following:

- **Strategy and direction:** Review the logframe (if one exists), focussing especially on the intended impacts. This can be followed by a brief Social Network Analysis exercise (e.g. carried out in a one-day workshop with staff), to assess current and potential relationships and how these may hinder or facilitate the intended impacts.
- **Management:** Carry out a ‘Lighter Touch’ Audit to assess which internal quality assurance and M&E procedures are already in place, and whether these are appropriate and sufficient. This information can be gathered through document review and staff interviews. Alternatively, a more discussion-based Horizontal Evaluation workshop might be held with staff to draw out strong and weak points of current internal organisational processes.
- **Outputs:** Depending on which outputs are considered most important by staff, and/or are most important in relation to achieving the intended impacts, the evaluation might include: (a) evaluation of a selected number of research reports (evaluating the quality of the research against appropriate academic criteria); and/or (b) evaluation of a selected number of policy briefing papers (perhaps using a focus group of policymakers to give feedback); and/or (c) evaluation of the website; and/or (d) evaluation of the knowledge network.
- **Uptake:** A brief exercise in Citation Analysis could be carried out, perhaps asking staff to choose a few of their prominent outputs and then drawing up an overview of how often and where these were cited (e.g. on other websites, in newspapers, etc).
- **Impact:** One or two more in-depth examples (e.g. Innovation Histories) that explore a time when the project, programme or institution did have an impact, how this was achieved, and how it might be replicated. This information may be gathered through document review, staff interviews, a workshop with staff, or a workshop with staff and invited external stakeholders.

Of course, the scope and depth of the evaluation will have to be adapted to the deadline, the budget available, and the amount of time that staff can give to the process. However, an attempt to include questions and reflection across more than one of the key performance areas may enable even a quick external evaluation to draw out core questions about impact, and thus to trigger reflection among staff on how their policy research project, programme or institution is making a difference.

References

- Acosta, A. and B. Douthwaite (2005) 'Appreciative inquiry: An approach for learning and change based on our own best practices', *ILAC Brief 6* (www.cgiar-ilac.org).
- Alwang, J. and V. Puhazhendhi (2002) 'The impact of the International Food Policy Research Institute's research program on rural finance policies for food security for the poor', *Impact Assessment Discussion Paper No 16*, International Food Policy Research Institute (IFPRI), Washington DC (www.ifpri.org/impact/iadp16.pdf).
- Arnold, E., J. Court, J. Stroyan and J. Young (2005) 'Evaluation of DFID's Engineering Knowledge and Research (EngKaR) Programme', ODI and Technopolis, (www.odi.org.uk/rapid/Projects/R0145).
- Barnard, G. (n.d.) 'The new frontier: Evaluation and organisational learning', Euforic (www.euforic.org).
- Barnett, A. (2006) 'Journeying from research to innovation: Lessons from the Department for International Development's Crop Post-Harvest Research Programme 'Partnerships for Innovation'', The Policy Practice (www.thepolicypractice.com/papers/1.pdf).
- Buchanan-Smith, M. (2003) 'How the Sphere Project Came into Being: A Case Study of Policy-making in the Humanitarian Aid Sector and the Relative Influence of Research', *ODI Working Paper 215*, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/Publications/Documents/WP215.pdf).
- Christiansen, K. with I. Hovland (2003) 'The PRSP Initiative: Multilateral Policy Change and the Relative Role of Research', *ODI Working Paper 216*, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/Publications/Documents/WP216.pdf).
- Butcher, C. and G. Yaron (2006) 'Scoping Study: Monitoring and Evaluation of Research Communications', background paper produced for the Communications M&E Group Workshop, 5-6 September 2006, London (available from Gabrielle Hurst: g.hurst@ids.ac.uk).
- Carrier, M., et al (2007) Evaluation Report of the Department of the Social Sciences and Humanities (D-GESS) at the ETH Zurich, Swiss Federal Institute of Technology (ETH), Zurich.
- CGIAR Secretariat (2003) 'Performance Measurement of Research Institutions and Research Programs – A Sourcebook', CGIAR Secretariat, The World Bank, Washington DC.
- CGIAR Working Group on Performance Measurement (2004) 'A plan for designing, developing, and implementing a performance measurement system for the CGIAR', CGIAR (www.cgiar.org/exco/exco6/exco6_wgpm_draft_imp_report.pdf).
- Coe, J., M. Luetchford and T. Kingham (2002) 'id21: Tracking routes towards impact', Institute for Development Studies (IDS), Brighton (www.ids.ac.uk).
- Cornwall, A. and G. Pratt (2004) 'Ideals in practice: Enquiring into participation in SIDA', *Lessons for Change in Policy and Organisations*, No 12, Institute for Development Studies (IDS), Brighton (www.ids.ac.uk).
- Creech, H. (2001) 'Measuring while you manage: Planning, monitoring and evaluating knowledge networks', International Institute for Sustainable Development (IISD) (www.dgroups.org/groups/pelican/docs/Creech2001_measuring_while_you_manage.pdf).
- Crewe, E. and J. Young (2002) 'Bridging Research and Policy: Context, Evidence and Links', *ODI Working Paper 173*, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/Publications/Documents/wp173.pdf).
- CTA (2006) *Strategic Plan 2007-2010*, ACP-EU Technical Centre for Agricultural and Rural Cooperation (CTA).
- Danish Ministry of Finance (2002) 'Improving an Organisation through Self-Assessment: The Common Assessment Framework (CAF)', presented at the Second Quality Conference for Public Administrations in the EU (www.zqconference.org/filer/404/CAF-english.pdf).
- David, R. and A. Mancini (2004) 'Going against the flow: Making organisational systems part of the solution rather than part of the problem. The case of ActionAid's Accountability, Learning and Planning System', *Lessons for Change in Policy and Organisations* No 8, Institute for Development Studies (IDS), Brighton (www.livelihoods.org/lessons/learning/GoingFlow.pdf).
- Davies, H., S. Nutley and I. Walter (2005) 'Assessing the impact of social science research: Conceptual, methodological and practical issues', Research Unit for Research Utilisation, University of St Andrews (www.st-andrews.ac.uk/~ruru).

- Davies, R. (2003) 'Network Perspectives in the Evaluation of Development Interventions: More than a Metaphor', paper presented at the EDAIS conference, 24-25 November (www.enterprise-impact.org.uk/pdf/Davies2.pdf).
- Davies, R. (2005) 'Moving from Logical to Network Frameworks: A modular matrix approach to representing and evaluating complex programs' (www.mande.co.uk/docs/MMA.htm).
- Davies, R. and J. Dart (2005) *The 'Most Significant Change' (MSC) Technique; A Guide to Its Use* (www.mande.co.uk/docs/MSCGuide.pdf).
- DFID (2006) 'Monitoring and Evaluation: A Guide for DFID-contracted Research Programmes', Central Research Department, DFID, London (www.dfid.gov.uk/research/me-guide-contracted-research.pdf).
- DFID (n.d.) *Social Development Toolkit*, UK Department for International Development (DFID), London (www.difd.gov.uk).
- Douthwaite, B. and J. Ashby (2005) 'Innovation histories: A method for learning from experience', *ILAC Brief 5* (www.cgiar.ilac.org).
- Douthwaite, B., S. Alvarez, S. Cook, R. Davies, P. George and J. Howell (2006) 'The use and potential of impact pathways in the Challenge Program on Water and Food (CPWF)', mimeo.
- Douthwaite, B. (2006) 'Making explicit practitioners' theories about how change happens', Presentation given at the Institute for Development Studies (IDS), Brighton, 11 July 2006 (www.ids.ac.uk/ids/aboutids/docs/practitioners_theories_for_change.ppt).
- Earl, S., F. Carden and T. Smutylo (2001) *Outcome Mapping; Building Learning and Reflection into Development Programs*, International Development Research Centre (IDRC) (www.idrc.ca/en/ev-9330-201-1-DO_TOPIC.html).
- ECDPM (2007) *ECDPM Strategy 2007-2011*, European Centre for Development Policy Management (ECDPM), Maastricht (www.ecdpm.org/Web_ECDPM/Web/Content/Navigation.nsf/index2?readform&www.ecdpm.org/Web_ECDPM/Web/Content/Content.nsf/7732def81ddd7a7ac1256c240034fe65/4c98d3dfoee95ef3c1257272002ed9ad?OpenDocument).
- ECDPM and ICEI (2005) 'EU mechanisms that promote policy coherence for development: A scoping study', European Centre for Development Policy Management (ECDPM), Maastricht, and Instituto Complutense de Estudios Internacionales (ICEI), Madrid (www.three-cs.net/3cs_publications).
- ETH (1999) 'Evaluation in Practice at ETH Zurich', Swiss Federal Institute of Technology (ETH) Zurich, mimeo.
- Gayfer, J. (2005) 'Country Programme Evaluations', Process Reflection Note, The Performance Assessment Centre (PARC) (www.parcinfo.org).
- Glanville, B. and Swanson, D. (2007) 'A Prototype Monitoring and Assessment System for Policy Influence at the International Institute for Sustainable Development', International Institute for Sustainable Development (IISD), mimeo.
- Haden, P. (2005) 'Quality Assurance: Lessons from other Organisations', The Performance Assessment Centre (PARC) (www.oecd.org/dataoecd/59/52/36502340.pdf).
- Hernes, G., Maxwell, S., Ndiritu, C., Rwegasira, D., Streefland, P., and Brylle, T. (2001) *Partnerships at the Leading Edge: A Danish Vision for Knowledge, Research and Development*, Report of the Commission on Development-Related Research Funded by Danida, Danida, Copenhagen (www.evaluation.dk).
- Hovland, I. (2005) *Successful Communication: A Toolkit for Researchers and Civil Society Organisations*, RAPID Toolkit, Overseas Development Institute (ODI), London (www.odi.org.uk/Rapid/Publications/Documents/Comms_toolkit.pdf).
- Hyatt, J. and A. Kaplan (2006) 'Seeing differently? Donors as learning organisations', Allavida (www.allavida.org/alliance/juno6e.html).
- IDRC Evaluation Unit (2005) 'Evaluation Strategy 2005-2010', International Development Research Centre (IDRC) (www.idrc.ca/uploads/user-S/11446799141Strategy_2005-10.pdf).
- IDRC Evaluation Unit (2006) 'Annual Corporate Evaluation Report 2006', International Development Research Centre (IDRC) (www.idrc.ca/uploads/user-S/11586128281ACE_2006.pdf).
- IDRC Evaluation Unit (n.d.) 'Reflections on evolving corporate performance assessment at the International Development Research Centre', mimeo.
- IISD (n.d.) 'IISD's Influencing Strategy', International Institute for Sustainable Development (IISD), mimeo.

- INTRAC (2006) 'INTRAC's 6th Evaluation Conference Report', International NGO Training and Research Centre (INTRAC) (www.intrac.org/pages/6thevalConference.html).
- ITAD in association with ODI (2000) *Evaluation of Danida's Bilateral Programme for the Enhancement of Research Capacity in Developing Countries (ENRECA)*, Danida Evaluation 2000/5, Danida, Copenhagen (www.evaluation.dk).
- Kassam, A.H., *et al* (n.d.) 'A framework for enhancing and guarding the relevance and quality of science: The case of the CGIAR', CGIAR Interim Science Council, FAO, Rome (<http://are.berkeley.edu/~sadoulet/papers/Relevance-QualofSci.pdf>).
- Keijzer, N. (2006) 'Mapping of approaches towards M&E of capacity and capacity development', ECDPM (www.dgroups.org/groups/pelican/docs/Mapping_M&E_capacity_o8o6o6.pdf).
- Leksmono, C., J. Young, N. Hooton, H. Muriuki and D. Romney (2006) 'Informal traders lock horns with the formal milk industry: The role of research in pro-poor dairy policy shift in Kenya', *ODI/ILRI Working Paper 266*, Overseas Development Institute, London (www.odi.org.uk/rapid/Publications/Documents/WP266.pdf).
- Lewison, G. (2005) 'Beyond SCI citations – New ways to evaluate research' *Current Science* 89 (9) 1524-1530.
- LTS International, Noragric and Oxford Policy Management (2005) 'Evaluation of DFID Renewable Natural Resources Research Strategy 1995-2005', EVD659, DFID, London (www.dfid.gov.uk).
- Matter, K., R. Mwai, E. Sefuke and A. Sherriff (2006) 'ECDPM External Evaluation 2001-2005' (www.gersterconsulting.ch/docs/Evaluation_ECDPM_report.pdf).
- McGann, J. (2006) 'Best Practices for Funding and Evaluating Think Tanks and Policy Research', prepared for the William and Flora Hewlett Foundation, mimeo.
- Patton, M.Q. (2002) 'Utilization-focused evaluation checklist' (www.wmich.edu/evalctr/checklists/ufe.pdf).
- Pestieau, C. (2003) 'Evaluating Policy Research', *Research Paper W 22*, Work Network, Canadian Policy Research Networks (CPRN), Ottawa (www.cprn.com/en/doc.cfm?doc=495).
- Pestieau, C. (2004) 'Post log to *Evaluating Policy Research: Can Outcome Mapping help in assessing the contribution of policy research?*' (www.idrc.ca/en/ev-67730-201-1-DO_TOPIC.html).
- Peterson, W., G. Gijbbers and M. Wilks (2003) *An Organisational Performance Assessment System for Agricultural Research Organisations: Concepts, Methods, and Procedures*, Research Management Guidelines No 7, International Service for National Agricultural Research (ISNAR) (www.isnar.cgiar.org/publications/rmg7.htm).
- QAA (2006) 'Handbook for institutional audit: England and Northern Ireland', Quality Assurance Agency for Higher Education (www.qaa.ac.uk/reviews/institutionalAudit/handbook2006/Handbook2006.pdf).
- Raitzer, D. and R. Lindner (2005) 'Review of the returns to ACIAR's bilateral R&D investments', *Impact Assessment Series Report No 35*, Australian Centre for International Agricultural Research (ACIAR), Canberra (www.aciar.gov.au/web.nsf/doc/ACIA-6G9VCN).
- Ramalingam, B. (2006) 'Tools for Knowledge and Learning: A guide for development and humanitarian organisations', RAPID Toolkit, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/publications/Documents/KM_toolkit_web.pdf).
- Rath, A., Björklund, G., Lansang, M.A., Saasa, O., and Sagasti, F. (2006) SAREC Support to International and Regional Thematic Research Programs, 2000–2005, Department for Evaluation and Internal Audit, SIDA Evaluation 06/40, Swedish International Development Agency (SIDA), Stockholm (www.sida.se/shared/jsp/download.jsp?f=Utvo6-40-1_SIDA31700en.pdf&a=26700).
- Ryan, J. (n.d.) 'State of the Art in Impact Assessment of Policy-Oriented Research in the CGIAR: A Scoping Study Report', mimeo.
- Ryan, J. (2003) 'Evaluating the impact of agricultural projection modelling using the IMPACT framework', *Impact Assessment Discussion Paper No 17*, International Food Policy Research Institute (IFPRI), Washington DC (www.ifpri.org/impact/iadp17.pdf).
- Ryan, J. and J. Garrett (2003) 'The Impact of Economic Policy Research: Lessons on Attribution and Evaluation From IFPRI', *Impact Assessment Discussion Paper No 20*, IFPRI, Washington DC (www.ifpri.org/impact/iadp20.pdf).

- Savedoff, W., R. Levine, N. Birdsall, *et al* (2006) *When Will We Ever Learn? Improving lives through impact evaluation*, Report of the Evaluation Gap Working Group, Centre for Global Development (CGD), Washington DC (www.cgdev.org/content/publications/detail/7973).
- Sayce, K., with P. Norrish (2006) 'Perceptions and Practice: An anthology of impact assessment experiences', Technical Centre for Agricultural and Rural Cooperation ACP-EU (CTA), Wageningen (www.anancy.net/index.php?specific=rss&file_id=12204).
- Schelhas, J. and L. Cervený (2002) *Social Network Analysis for Collaboration in Natural Resource Management* (www.partnershipresourcecenter.org/resources/publications/social_network_analysis.php).
- Solesbury, W. (2003) 'Sustainable Livelihoods: A Case Study of the Evolution of DFID Policy', *ODI Working Paper 217*, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/Publications/Documents/WP217.pdf).
- Taylor, J. (2001) 'Evaluating websites', International NGO Training and Research Centre (INTRAC), Oxford (www.intrac.org/resources_database.php?id=49).
- Thiele, G., A. Devaux, C. Velasco and K. Manrique (2006) 'Horizontal evaluation: Stimulating social learning among peers', *ILAC Brief 13* (www.cgiar-ilac.org/downloads/Briefs/ILAC_Brief13.pdf).
- UNDP (1998) *POET User's Manual: Participatory Organisational Evaluation Tool*, United Nations Development Programme (UNDP) (www.cominit.com/pdf/POET_UsersManual.pdf).
- Watts, J. (2005) 'Learning-oriented evaluation: A tool for promoting institutional learning and program improvement', *ILAC Brief 3* (www.cgiar-ilac.org).
- Weiss, C. (1977) 'Research for Policy's Sake: The Enlightenment Function of Social Research', *Policy Analysis 3* (4) 531-545.
- White, H. (n.d.) 'The road to nowhere? Results based management in international cooperation', Euforic (www.euforic.org).
- Wilson, F. and Brown, L.D. (2004) 'The Development Research Centre on Participation, Citizenship and Accountability: Mid Term Review – Institutional Development', Citizenship DRC, Institute for Development Studies (IDS), Brighton.
- World Bank (1996) 'Performance Monitoring Indicators: A Handbook for Task Managers', Operations Policy Department, World Bank, Washington DC (www.worldbank.org/html/opr/pmi/pmi.pdf).
- World Bank (2004) 'Monitoring and Evaluation; Some Tools, Methods and Approaches', World Bank, Washington DC (www.worldbank.org/ieg/ecd/tools/).
- Young, E. and L. Quinn (2002) 'Writing Effective Public Policy Papers: A Guide to Policy Advisers in Central and Eastern Europe', *LGI Documents*, (<http://lgi.osi.hu/publications/default.asp?id=112>).
- Young, J., J. Kajume and J. Wanyama (2003) 'Animal Health Care in Kenya: The Road to Community-Based Animal Health Service Delivery', *ODI Working Paper 214*, Overseas Development Institute (ODI), London (www.odi.org.uk/rapid/Publications/Documents/WP214.pdf).

Appendix 1. Examples of institutional evaluations of policy research institutes and think tanks

CGD

The Center for Global Development (CGD) in Washington DC was recently evaluated by a group of five experts. The evaluation was commissioned by CGD's funders (the Bill and Melinda Gates Foundation, The William and Flora Hewlett Foundation, The John D. and Catherine T. MacArthur Foundation, and the Rockefeller Foundation), and covered CGD's first five years of operations. The evaluation focused on the following core questions:

- Does CGD's *research agenda* fill a development policy niche and meet the needs of its policymaking targets?
- Does CGD's *research product* stand up to the highest standards of technical rigor and influence the work of other leading researchers?
- Is CGD's *communications and outreach strategy* achieving its desired impact and is it appropriately institutionalized?
- Has CGD built and leveraged appropriate *partnerships*?

The evaluation review consisted of over 150 interviews, survey responses from more than 1,400 end-users, a competitive mapping of the think tank community, and independent white papers covering broad trends in think tank development. The report also included in-depth descriptions and assessments of three selected case studies of CGD's work: the Making Markets for Vaccines Initiative, the Nigerian Debt Relief Initiative, and the Population Dynamics and Economic Development Initiative. In addition, the evaluation included a thorough analysis of CGD's organisational management systems and governance. For the purpose of confidentiality, this section was removed from the public version of the report.

The evaluation team concluded that CGD has carved an appropriate niche for itself in the think tank community, that their research initiatives have demonstrable impact, they they need to engage more broadly with policymakers in rich countries as well as with advocacy groups and organisations with a variety of perspectives, and that the CGD leadership should nurture deeper relationships with numerous key stakeholders.

Relevant documents:

- Bumgarner, R., Hattaway, D., Lamb, G., McGann, J., and Wise, H. (2006) *Center for Global Development: Evaluation of Impact*, Arabella Analytical Report, Arabella Philanthropic Investment Advisors (www.arabellaadvisors.com/CGDreport/CGDWeb.pdf).

CGIAR

The Consultative Group on International Agricultural Research (CGIAR) consists of a network (a 'system') of affiliated institutions across the globe. CGIAR system-wide M&E therefore goes beyond the institutional level to include the entire system. This section presents their system-wide M&E plan. For the evaluation of an individual CGIAR institution, see the section below on IFPRI.

CGIAR produced a sourcebook on performance measurement of research institutions and programmes in 2003 (CGIAR Secretariat, 2003) and a planning document the following year (CGIAR Working Group on Performance Measurement 2004). The Sourcebook found that the main assessment methods of research performance used were bibliometric analysis and peer review. The CGIAR Working Group on Performance Measurement did not wish to adopt any of these models wholesale, but rather chose

elements to design their own CGIAR system-wide **Performance Measurement System**. The selected core elements are (CGIAR Working Group on Performance Measurement, 2004: 8):

Results

- Outputs
- Outcomes
- Impacts
- Stakeholder perceptions

Potential to perform

- Quality and relevance of staff
- Quality and relevance of programmes
- Governance and institutional health
- Financial health

CGIAR centres are now asked to report on indicators for each of these elements annually. The Performance Measurement System complements other existing M&E processes, such as project reviews, centre-managed reviews, and centre-commissioned external reviews.

Relevant documents:

- CGIAR Secretariat (2003) 'Performance Measurement of Research Institutions and Research Programs – A Sourcebook', CGIAR Secretariat, The World Bank, Washington DC.
- CGIAR Working Group on Performance Measurement (2004) 'A plan for designing, developing, and implementing a performance measurement system for the CGIAR', CGIAR (www.cgiar.org/exco/exco6/exco6_wgpm_draft_imp_report.pdf).

CHSRF

The Canadian Health Services Research Foundation (CHSRF) in Ottawa has recently undergone its second five-year institutional evaluation. The first international review was completed in 2002. The second has been carried out 2006–7. The evaluation consisted of internal evaluation by staff, collected in three background 'briefing books':

- The first of these books (36 pages) covers the foundation's history, the panel's terms of reference, the evaluation philosophy (focusing on the institution's strategic objectives and desired impacts), and a summary of the report from the 2002 International Review Panel, along with the foundation's responses to its recommendations.
- The second briefing book (63 pages) is a description of programmes and activities (the *inputs*), the resulting products and deliverables (the *outputs*) as well as data on *programme-specific outcomes*, organized by strategic objective.
- The third briefing book (76 pages) is evaluative data for the *overall impact* of all the foundation's programmes on the cultures of research and decision-making in the health sector – the synergistic effect of activity under all four of the strategic objectives. This is focused on the results of survey work designed to assess this overall impact supplemented with boxed illustrations of such impact.

An external contractor was also commissioned to carry out a consultation exercise on stakeholder perceptions, which was presented in a fourth briefing book.

- The fourth briefing book (61 pages) is an appraisal of how the foundation's stakeholders view it and its work. It contains a summary of the responses received to an open-ended call for comment to stakeholders.

These briefing books were then considered by an international external panel. The panel was made up of four CEOs/Directors of health policy research centres in Lisbon, Washington DC, London, and Alberta. They received the material for evaluation from October-December 2006, and submitted their final report at the end of March 2007.

Relevant documents:

- Report highlights, the Final Report, and CHSRF's response to the international panel's comments, are all available at www.chsrf.ca/about/ga_accountability_impact_ol_e.php

Citizenship DRC

The DFID-funded Development Research Centre (DRC) on Participation, Citizenship and Accountability, hosted by the Institute for Development Studies (IDS) in Brighton, UK, was evaluated as part of a mid-term review in 2004. This has been included as an example of a traditional **DFID-funded mid-term review**. The review focused on the achievements of the Citizenship DRC in terms of its original objectives and its potential to impact on policy. The review team consisted of Fiona Wilson, from the Danish Institute for International Studies who focused on assessing the research products of the DRC, and L. David Brown, from the Hauser Centre for Nonprofits Organisations at Harvard University, who assessed its institutional development, capacity building, and dissemination and policy influence aspects.

The review drew on information from several sources: Wilson read the research output of the programme to get an overview of its nature, held a brain-storming session with the Co-ordination Team, conveners and members present at IDS at the time and then made a more detailed assessment of the primary products of four streams of work. Brown reviewed a variety of archival materials, interviewed twenty-one participants, and facilitated two days of reflection with DRC participants on its impacts and its future at a workshop in Brazil. Then both reviewers met for several days at IDS with the DRC Coordination Team and DFID staff to review and extend initial analyses and discuss the recommendations summarised in their report.

The report highlighted that the Citizenship DRC has chosen to develop a **partnership form** rather than the more common centralised form of development research centre. The evaluation team concluded that the DRC has been quite successful in building a mutually influential partnership for collaborative inquiry, but that doing so has required considerable investment of time, resources, and funds. They also concluded that the Citizenship DRC has produced several streams of important research, though those streams were unevenly developed at the time (in 2004), and recommendations were given as to how the DRC could complete and expand the relevance of those streams most advanced at that point. While the evaluation team did not have information to assess policy impacts directly, the available evidence indicated that the DRC had considerable potential to influence the knowledge contexts within which policies are made, and in some countries their partners have positions within policymaking networks that grant them significant influence.

Relevant documents:

- Wilson, F. and Brown, L.D. (2004) 'The Development Research Centre on Participation, Citizenship and Accountability: Mid Term Review – Institutional Development', Citizenship DRC, Institute for Development Studies (IDS).

CTA

The ACP-EU Technical Centre for Agricultural and Rural Cooperation (CTA) in the Netherlands has recently undergone an external evaluation. CTA's Strategic Plan/Framework for Action 2001–2005 was

evaluated by an external consultancy team between January and August 2005. The aim of the evaluation was twofold:

- To analyse the implementation of CTA's mission in terms of the goals set in the plan,
- To obtain baseline information that would serve as inputs to the next strategic plan (2007–2010).

After completion of the evaluation report in January 2006, CTA began preparing the current strategic plan in February 2006 through participatory processes involving staff members, ACP and EU stakeholders and partners. The drafting of plan was completed in August 2006, and reviewed between September and November 2006.

Relevant documents:

- CTA (2006) *Strategic Plan 2007-2010*, ACP-EU Technical Centre for Agricultural and Rural Cooperation (CTA).

ECDPM

The European Centre for Development Policy Management (ECDPM) in Maastricht recently commissioned **an external evaluation visit** to evaluate the work of the institution as a whole (Matter, Mwai, Sefuke and Sherriff 2006). ECDPM draws up its institutional plans for five-year cycles, and so the evaluation covered the last five-year cycle from 2001-2005.

The evaluation visit had four specific objectives:

- To assess the pertinence of the Centre as an independent foundation, taking into account its mandate, strategic and methodological choices, networks, partnerships and stakeholders;
- To evaluate the effectiveness of the Centre's positioning, external operations and networks;
- To identify and trace plausible patterns of Centre impact regarding the policy processes it has been directly involved in, emphasising both strong and weak points; and
- To formulate recommendations for the further institutional development of the Centre, providing feasible options and future scenarios for consolidation and improvement.

The evaluation was undertaken from February to May 2006 by a four-person team. Since ECDPM works to build an effective partnership between the European Union (EU) and Africa, the Caribbean and Pacific (ACP), the evaluation team was composed of two experts from Europe (including the team leader) and two from the ACP (both from Africa). The team used two evaluation approaches: (1) document review, including internal and external documents, and a self-assessment prepared by ECDPM; and (2) semi-structured interviews with a sample of ECDPM staff, Board members, and representatives of external stakeholders across the globe.

The final evaluation report addresses the following topics: ECDPM's mandate in light of the changing context; ECDPM's process approach; ECDPM's partnership and network approach; ECDPM as an independent broker; Pertinence and performance of programmatic choices, including the programmes on 'ACP-EU trade relations', 'Political dimensions of partnerships', 'Actors of partnerships', and 'Internal donor reform'; Information and (external) communication; Institutional management and development; Finance and fund-raising; Added value; and Impact.

Relevant documents:

- Matter, K., R. Mwai, E. Sefuke and A. Sherriff (2006) 'ECDPM External Evaluation 2001-2005' (www.gersterconsulting.ch/docs/Evaluation_ECDPM_report.pdf).
- ECDPM (2007) *ECDPM Strategy 2007-2011*, European Centre for Development Policy Management (ECDPM), Maastricht (www.ecdpm.org/Web_ECDPM/Web/Content/Navigation.nsf/index2?readform&http://www.ecdpm.org/Web_ECDPM/Web/Content/Content).

nsf/7732def81dddafa7ac1256c240034fe65/4c98d3dfoee95ef3c1257272002ed9ad?OpenDocument).

IDRC

The International Development Research Centre (IDRC) in Canada has been one of the pioneers in the field of evaluating policy research, mainly through their development of the Outcome Mapping approach (Earl, Carden and Smytulo, 2001).

When it comes to its own institutional assessment, IDRC uses a **decentralised evaluation system** (IDRC Evaluation Unit, 2005; 2006) with a utilisation focus, i.e. it is set up to involve programme staff so that the results and lessons will be used in future work (Patton 2002). The evaluation system relates to IDRC's Corporate Assessment Framework (CAF), which outlines seven Performance Areas for the institution: enhancing research capacities, research results for policy and technology influence, collaborating with Canadians, strategic knowledge gathering, gender equality and women's rights, donor partnerships, and evaluative thinking. IDRC's decentralised evaluation system consists of the following elements:

- *Programme and project evaluation reports* from IDRC's eighteen programme initiatives (though programmes are not required to report every year). Each report addresses the Performance Areas that are relevant to the programme or project. Since the evaluation system is user-focused, the programmes and projects decide on evaluation questions themselves.
- An *Annual Corporate Evaluation Report*, written by the IDRC Evaluation Unit. This combines findings from the programme and project reports in a concise way (the 2005-06 report was 39 pages including annexes) (IDRC Evaluation Unit 2006). The annual report synthesises the results from the programme and project reports, and then focuses specifically on a few selected topics, themes and particularly interesting projects.
- *Reviews by external experts*. Programmes initiate external reviews. IDRC have noted that while external reviews can provide useful external perspectives, they also require considerable staff time and effort.
- Other reports, including *Regional Director Reports* (every two years), *Director of Program Area Reports* (every two years), and *Rolling Project Completion Reports*

Relevant documents:

- IDRC Evaluation Unit (2005) 'Evaluation Strategy 2005-2010', International Development Research Centre (IDRC) (www.idrc.ca/uploads/user-S/11446799141Strategy_2005-10.pdf).
- IDRC Evaluation Unit (2006) 'Annual Corporate Evaluation Report 2006', International Development Research Centre (IDRC) (www.idrc.ca/uploads/user-S/11586128281ACE_2006.pdf).

IFPRI

The CGIAR-affiliated International Food Policy Research Institute (IFPRI) in Washington DC has been working on impact assessment of its policy research since the mid-1990s. They have found that a **case study approach** is the most appropriate to articulate, measure and document the impacts of their policy research. Beginning in 1998, they undertook a series of case studies that examined relevant policy processes and the use of information by policymakers. They set up an *Impact Assessment Discussion Paper* series to publish their studies (www.ifpri.org/impact/impact.htm). These case studies are similar to the Innovation Histories mentioned above. They identified a policy research project and then traced the stages it went through and the impact it had on its surroundings.

Ryan and Garrett (2003) chart the subsequent developments in IFPRI: After reviewing the lessons learned from the initial case studies, the Board of Trustees in 2000 requested that management prepare an operational strategy to institutionalise impact evaluation at IFPRI. A Working Group on Impact Evaluation (WGIE) was established. One of their responsibilities was to ensure that impact evaluation was integrated in the strategy; another was to instil a culture of impact evaluation within IFPRI. For this purpose, all four division directors and five IFPRI staff were members, along with three non-IFPRI members. In 2001, a number of pilot exercises involving *ex ante* impact evaluation of new projects was tried out as part of the new strategy, and IFPRI also conducted evaluations of some of its thematic research programmes (Alwang and Puhazhendhi, 2002; Ryan, 2003).

IFPRI still found that the evaluations remained at a certain remove from the daily operations of staff (Ryan and Garrett, 2003). Steps were taken to address this so that the evaluations could become real learning opportunities. In 2002-03, for example, all research staff were requested to narrate instances where their research outputs had influenced policy and had subsequent social or economic impacts. These were conducted in focus groups of four to six staff from the different research divisions in order to stimulate cross-fertilisation. Since then IFPRI has continued its focus on impact assessment through further publications and internal workshops.

Relevant documents:

- Ryan, J. and J. Garrett (2003) 'The Impact of Economic Policy Research: Lessons on Attribution and Evaluation from IFPRI', Impact Assessment Discussion Paper No 20, IFPRI, Washington DC (www.ifpri.org/impact/iadp20.pdf).

IISD

The International Institute for Sustainable Development (IISD), in Winnipeg, Canada, first drew up an institutional Influencing Strategy, and then developed an institutional Monitoring and Assessment (M&A) framework based on the Influencing Strategy.


The **Influencing Strategy** broadly sets out principles for IISD's work. The strategy recognises that IISD's work should always involve other key stakeholders that they need to engage with in order to achieve their aims. The strategy then points to a number of questions that need to be asked across all work initiatives, including how to maintain important connections over time (Relationship Management), what the knowledge needs of IISD and other key stakeholders are (Knowledge Management), and how to take advantage of policy windows and how to create opportunities (Opportunity Management).

The **M&A framework** draws its main questions from the Influencing Strategy, and presents these in a **prototype self-assessment report card** (see Figure 18 below). This method was run as a pilot test in 2006, and was broadly considered to be a success. IISD found that the basic framework they were using was robust and that it had helped to sharpen programme goal statements. They found that data for the majority of indicators could be obtained from existing IISD reporting processes, and in particular that data from anecdotal accounts of project staff can yield insights on IISD's own perception of influence on policy and behavior change. They also learnt that they do not have sufficiently detailed and verifiable data for the key indicators linked to policy and behaviour change, and that at the moment data is not systematically collected after the end of projects. These lessons are now being used to simplify and improve the indicator set for data collection in 2007 and to refine data collection methods for policy and behavior change at the programme level.

Relevant documents:

- IISD (n.d.) 'IISD's Influencing Strategy', IISD, memo.
- Glanville, B. and Swanson, D. (2007) 'A Prototype Monitoring and Assessment System for Policy Influence at the International Institute for Sustainable Development', IISD, memo.

Figure 18. IISD Policy Influence Report Card

 2005-06 IISD Policy Influence Report		Learning and Improvement		
Framework for Effective Policy Influence		Learning Analysis	Improvement Options	
Category	Aspect	Indicator	Source	
Policy and Process Changes Are we seeing progress toward the policy decisions or process changes we wish to influence?	Policy Influences	Number and nature of recommended changes for sustainable development realized for each strategic direction	Project Lessons Learned Database	
	Project Diversity	Number of projects implemented which could impact each strategic direction	Project Lessons Learned Database	
Effective Relationship Management Are we effectively managing relationships with those people in a position to influence the type of changes we seek?	Behavioral Influences	Total number and nature of behavioral changes we may have influenced in decision makers, learners, others.	Project Lessons Learned Database	
	Repeat Funding	Number and type of new funders	Annual report	
	Program Influence Strategies	% of funders which are repeat funders % of program goals with an influencing strategy	Annual report Knowledge Communications program statistics	
Effective Knowledge and Opportunity Management Are we effectively creating and communicating knowledge needed by those in a position to influence change? Are we able to leverage opportunities to communicate with decision makers?	Nature of our Work	Number and type of knowledge outputs developed (e.g. policy frameworks, case studies, indicators, etc.)	Project Lessons Learned Database	
	Innovation	Number of ideas incubated Number of proposals submitted to innovation fund per year and per competition	IISD Sustainability Report Innovation Fund data	
	Communication and Engagement	Number of publications generated Proximity to decision maker	Nature of innovation in projects implemented	Project Lessons Learned Database Information Centre data
		Number of times IISD appears in the news per year	Total number of document downloads per year	Project Lessons Learned Database Knowledge Communications program statistics
Diversity in Human Capital	Total number of email lists and email list subscribers	Total number of nationalities represented among staff & associates	Knowledge Communications program statistics IISD Sustainability Report	
	Number of different academic degrees among staff & associates (based on highest degree obtained)	Female / Male ratios	IISD Sustainability Report IISD Sustainability Report	

Source: Glanville and Swanson (2007: 3)

ODI

The Overseas Development Institute (ODI) in London is undergoing a Strategic Review in 2007, and one of the outcomes of this Strategic Review will be an institutional M&E Strategy that will guide monitoring and evaluation activities for the next organisational phase, with a particular focus on ongoing monitoring. The steps leading up to the M&E Strategy include the following:

- The present *Survey Paper* ('Making a Difference: M&E of Policy Research', by Ingie Hovland);
- The establishment of an M&E Working Group;
- Internal workshop to 'launch' the process, to inform staff, and to invite participation from those especially interested;
- Collect and review programme evaluations from across the organisation (*Paper 1: Review of Existing Evaluations*);
- Draw up an audit/inventory of M&E approaches that are currently in use across the organisation (*Paper 2: Audit of Existing M&E Systems*);
- Collect around ten brief stories of change from ODI's work, asking relevant staff to write them up, and drawing out comparisons in a half-day workshop (*Paper 3: Stories of Change*);
- Peer review session with a small number of invited external experts, to gather feedback and comments in an informal and constructive manner;
- An internal meeting to present and discuss the final draft of the institutional *M&E Strategy for ODI*.