

## Chapter 2 Evaluation Methodology and Criteria

### 2.1 Preliminary Activities by PARC

2.1.1 In 1997, DFID commissioned a study to examine the monitoring of the impact of the RNRRS<sup>2</sup>, which concluded that there was no case for undertaking impact assessments as a routine exercise, but that there could be a case for commissioning assessments of the impact of project clusters, which had the potential to affect large numbers of people. A more recent study by Flint and Underwood (2002) indicated that there was likely to be difficulty in assessing the impact of the RNRRS on the poor, because of a lack of common indicators of impact amongst the programmes, the lack of a systematic approach to monitoring and evaluation, and particularly of programme uptake. DFID's Central Research Department (CRD) then commissioned the Performance Assessment Resource Centre (PARC) in 2003 to establish commonly agreed benchmarking tools to enable impact assessment across the ten research programmes.

2.1.2 Four tools were developed:

- Structured Impact Matrix (SIM) to capture a comprehensive spreadsheet of all the projects undertaken through RNRRS funding, with details to indicate focus and achievement (for all projects) and the availability of qualitative and quantitative data against which to assess project achievement;
- Impact Pathways (IP) for current sets of projects to demonstrate how particular projects or clusters of projects were designed to meet the overall purpose of the programme;
- Impact Assessment Questionnaires, for current projects which often built on earlier projects, to provide details about each project's focus and achievement, in order to show evaluators where they might search for further relevant detail; and
- Impact Timelines to show how individual projects or clusters of projects actually progressed towards the programme purpose over time.

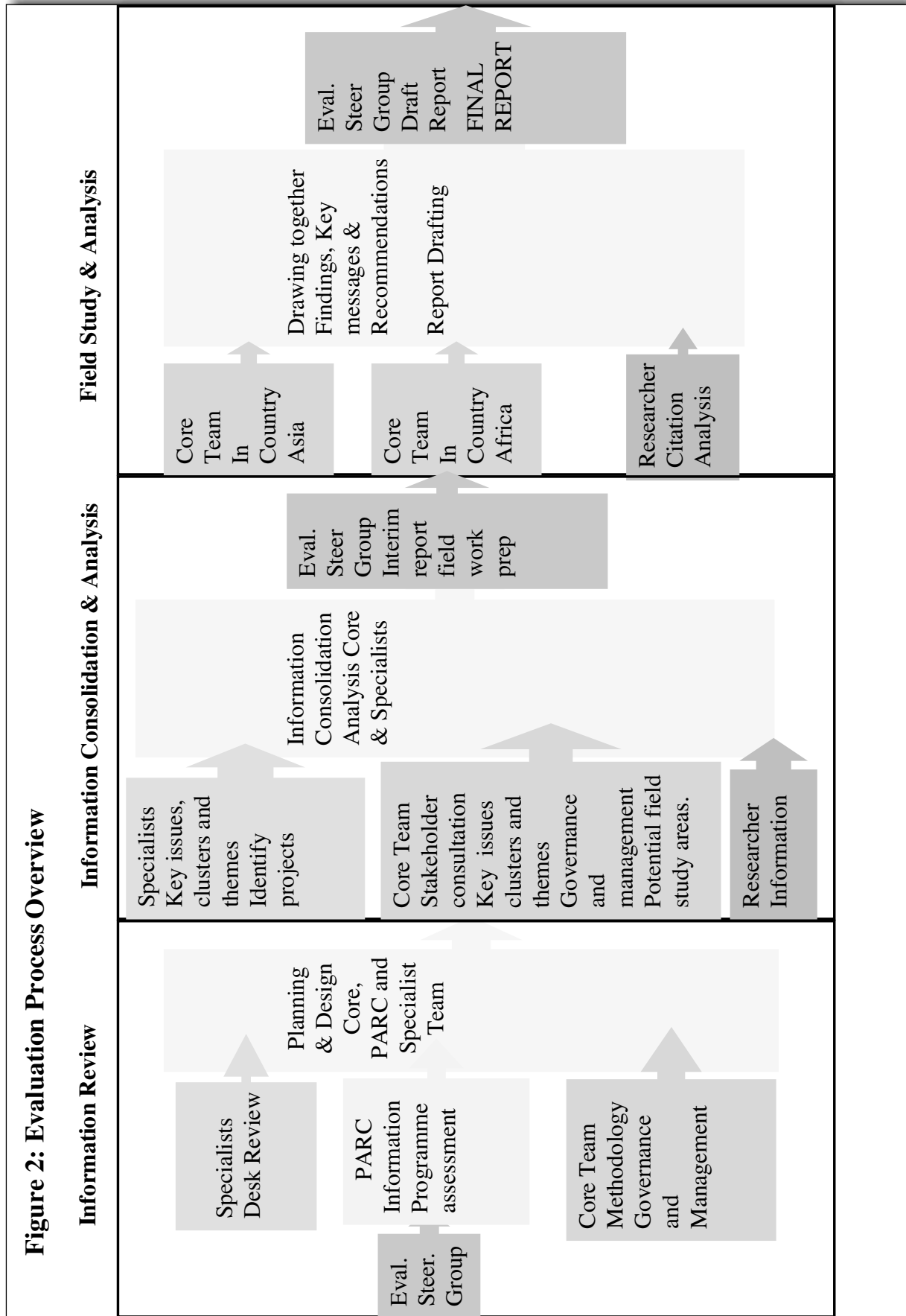
2.1.3 Programmes applied the different tools to different degrees with only three providing some data using the Impact Assessment Questionnaire and Impact Timelines. Furthermore, most of what was expected to be background information became available after the evaluation had commenced, making it necessary for the Core Team to launch their information gathering and consultation in parallel. Nonetheless, the PARC (2004) report provided information on the timelines of the RNRRS, a summary of governance arrangements, the structures of the programmes, and the impact pathways anticipated by the PMs.

### 2.2 The “stop-check-go” staged process

2.2.1 A “stop-check-go” staged process was adopted for the evaluation (Figure 2). It enabled the Core Team to draw together, collate and analyse the vast amounts of information provided by the RNRRS programmes. Each stage allowed engagement with different stakeholder groups and therefore provided for cross-validation and triangulation of findings in previous stages. Work plans were reviewed and refined as required at the end of each stage. An important factor in adopting the methodology was to ensure the most efficient and effective use of the personnel and resources, including those of the PARC.

---

<sup>2</sup> Don Brown et al. (1997). Monitoring the Impact of the DFID Renewable Natural Resources Research Strategy for 1995-2005, ITAD for the NRRD.



- 2.2.2 To make the most efficient and effective use of the experience and skills of the Core Evaluation Team the RNRRS Strategy was assessed under three key facets: quality of science, impact or potential impact on poverty, and quality of management. The process of breaking down the Strategy into these key facets allowed the evaluation team to give an independent assessment of the primary driving forces behind the Strategy. The team could then derive key lessons, identify best practices and key influences and the balance relating management, and good science and impact on poverty when deriving recommendations for future orientation of the Strategy.
- 2.2.3 The key facet approach enabled the evaluation team to break down standard internal RNRRS reporting against the ‘nested’ logical frameworks which primarily address the overlap between the key facets, see Figure 3.

**Figure 3 Key Elements of the Evaluation**



### **2.3 RNRRS and Individual Programmes**

- 2.3.1 The RNRRS has been operating for 10 years 1995 -2005. The Strategy is now comprised of ten, individually contracted research programmes which collectively amount to over 1600 research projects. The amount of information that has been associated with the RNRRS is enormous. Table 1 shows the distribution of budgets and projects by programme through the whole lifespan of the RNRRS.

**Table 1 Budgets and projects of RNRRS programmes, 1995 - 2004**

Programme	Total Budget (million pounds) <sup>a</sup>	% Total budget	% Budget proposed in Yellow Brick	Number of projects <sup>b</sup>			
				Latin America	Asia	Africa	Total
AHP	22.38	11.84	8.1	8	21	81	110
Aquaculture	2.66	1.41	2.0				
Fish Genetics	2.69	1.42	2.0				
AFGP	3.79	2.01		0	48	5	53
LPP	13.23	7.00	8.4	19	21	78	118
CPHP	23.41	12.39	12.6	5	32	173	210
CPP	49.26	26.07	26.4	45	130	310	485
FMSP	5.11	2.70	2.0	2	30	13	45
FRP	22.07	11.68	13.2	73	57	86	216
NRSP	27.21	14.40	14.2	28	74	95	197
PSP	13.54	7.17	8.4	8	93	58	159
PHFRP	3.62	1.92	1.5	0	13	13	26
Flexible			1.2				
<b>Total</b>	<b>188.97</b>	<b>100.00</b>	<b>100.00</b>	<b>188</b>	<b>519</b>	<b>912</b>	<b>1,619</b>

a Source PARC, 2004, Figure 9. Original data from Elizabeth Warham, DFID - No data for 1995/96 for FRP and PSP, no data for 1996/97 for LPP, CPHP, PRSP, no data for FRP for 1997/98, no data for LPP and CPP for 1998/99.

b Source: PARC, 2004. Original data from Elizabeth Warham, DFID. CPHP has noted a discrepancy between their figures and the PARC data.

## 2.4 Sampling and Data Collection

2.4.1 The Core Team used a number of methods and techniques for collecting the information needed for the evaluation. The total population of Programme Managers (PMs), Programme Advisory Committee (PAC) Chairpersons, and DFID Lead Advisers (LAs) were interviewed, sometimes more than once, using specific protocols (Annex 5). In consultation with DFID and PMs the UK based institutions which have led the most projects were interviewed, also using a specific protocol (see Annex 5 for protocols and Annex 6 for people interviewed).

2.4.2 The Evaluation Team has also assimilated enormous amounts of information from documented sources including:

- Strategic management and policy documents.
- Guidelines for operational and contractual arrangements.
- Standard reporting formats for programmes and projects.
- Strategic reviews for the RNRRS and for the individual programmes.
- Output reports and publications from projects and programmes.

- 2.4.3 For detailed examination of the scientific quality from the programmes a purposive sample of past and current projects were selected for detailed study. With over 1600 DFID-registered projects spread over 11 (later 10) programmes<sup>3</sup>, it was not feasible to cover the whole population given the time and resources of the evaluation. Furthermore, given the incomplete and sometimes outdated information, and differing coverage of Programmes in both the DFID database and the PARC SIMs, it was not thought wise to select a random sample of projects for detailed analysis. Purposive samples of projects were selected by Specialists in close collaboration with PMs to represent closed and current projects, large and small projects, the span on the A-H Scale<sup>4</sup>, and projects considered by PMs as the most and least successful. Two to eight percent of all projects ever funded were selected from each programme.
- 2.4.4 The individual Specialists' approach to sampling had to be purposive and be taken through a process of iteration. To ensure independence, it was required that the Specialist Reviewers had not been directly engaged in the programme they were evaluating during the period from 1994. Individual Specialists were, of course, aware to different degrees of the research that had been undertaken, of the more important findings and of the programmes themselves. This is inevitable given the high profile of the research funded, especially in those partner countries where substantial research has been supported by RNRRS, and the efforts made by Programme Managers to publicise their programme achievements and outcomes.
- 2.4.5 Each specialist adopted an approach to sampling that differed in detail depending on the characteristics of the programme. The aim was not to evaluate individual projects. It was to secure an understanding of a wide range of projects, that represented varying degrees of 'success' and to use these as proxies and examples to answer the main questions on the quality of the research, its management and delivery and on the use made of the results. (Section 2.5).
- 2.4.6 All individual programmes contained examples of projects that had failed for external reasons, which nevertheless assisted the team in identifying how programmes had extracted and adopted relevant lessons. Within the range of variation that applies in each individual programme, the Specialist Reviewers used a combination of reading, interrogation and discussion to secure a sample of projects for further detailed study which covered the range of basic and applied science, one-off and sequential projects, different contractors and so on.
- 2.4.7 For in-country stakeholder consultations, the four countries with the highest number of projects in each RNRRS region were identified (Bolivia, Costa Rica, Honduras and Mexico in Latin America, Bangladesh, India, Indonesia and Nepal in Asia, and Ghana, Kenya, Tanzania and Uganda in Africa). The number of projects in each country ranged from 20 to 144. Because of time constraint and logistics, and the low number of projects (20-49 per country), stakeholder consultations were not held in Latin America, as well as in Nepal and Tanzania.

<sup>3</sup> Fish genetics merged with aquaculture under the management of Stirling University as the University of Wales did not want to renew its contract due to Welsh legislation which precluded research into genetic modification of organisms

<sup>4</sup> All DFID-funded projects are scored with Poverty Aim Markers to show where they intend to focus in terms of the continuum of effort involved in poverty reduction. The Key to Scoring for Uptake Pathway are as follows:  
 A – Generation of relevant research results  
 B – Formal/informal agreement with target institutions  
 C – Development of appropriate research-based products through adaptation/packaging  
 D – Promotion of products into target institutions  
 E – Adoption of products by target institutions  
 F – Application and replication of results in target institution programmes  
 G – Promotion of technology or behavioural change among end-users by target institutions  
 H – Adoption of technology by end-users and generation of economic benefits

Projects were also categorized as follows:

- an  poor people e.g. influencing policy change to permit the use of genetically modified crops;
- an “inclusive” aim indicates that wider social groupings, including poor people, may also benefit from the project - in the context of RNRRS an inclusive focus also encompasses influencing institutional change;
- a “focused” aim indicates that the primary aim of the project is to ensure a more exclusive focus on benefits to specific groups of poor people.

2.4.8 The Core Team split into two groups and visited Bangladesh, India and Indonesia in Asia, and Ghana, Kenya and Uganda in Africa. DFID painstakingly constructed a list of project participants in each of the countries and the Core Team interviewed as many as could be contacted, either individually or in groups of up to 12 participants in each country using a structured set of questions (Annex 5). Stakeholders in about 10% of all RNRRS projects ever funded were interviewed (Table 2 and Annex 6). They included farmer organisations, NGOs, NARS, IARCs, private sector, extension agents, Ministries, and DFID officials.

**Table 2 Number of Projects Sampled During in-Country Visits, by Country and Programme**

Country	LPP	CPP	CPHP	NRSP	PSP	FRP	AFGP	PHFRP	FMSP	AHP
Indonesia	2	1	2	1		3			1	6
Bangladesh	4	6	5	13			1		3	
India	5	4	5	3	13	2	2			
ASIA	11	11	12	17	13	5	3	0	4	6
Ghana	2	10	14	5	3	2				
Kenya	4	10								1
Uganda	5	16	6	2						0
AFRICA	11	36	20	7	3	2	0	0	0	1
TOTAL	22	47	32	24	16	7	3	0	4	7

## 2.5 Criteria for Evaluation of Science Quality

2.5.1 Science quality was the main target of the specialist evaluation. Impacts on the poor, as well as management and governance aspects are covered mainly by the Core Team. However, in the course of this more detailed science evaluation, Specialists were requested to form impressions of the degree to which the projects, and through them the programme, had been conducted in a participatory mode, and to pass the impressions on to the Core Team.

2.5.2 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. Since the methodology including biometrics, research process and reporting of results is subject to review by qualified peers, they provide an independent assessment of the quality of the research and the validity of the output. An assessment of a sample of the publications from RNRRS programmes, including citation analysis, was undertaken by the Evaluation Researcher.

2.5.3 Specialists used the criteria and factors below to score the programmes. The Scores were used as a mechanism for summarising the overall assessment of the programmes, but weighted averages across programmes were not intended to be, and were not used by the Core Team to arrive at judgements for all of RNRRS as that would have involved too much value-judgement about the variety of projects in the programme portfolios, in view of the limited and purposive nature of the samples. Rather, The Core Teams used the results, in addition to its information derived from its own interviews of PMs, project leaders and in-country stakeholders to arrive at broad conclusions on the quality of scientific output.

2.5.4 The methodology for the citation analysis can be found in Annex 11. The citation analysis undertaken within the evaluation is only for use as a guide. It was never the intention to compare the individual programmes within the RNRRS with each other. The sampling methodology for published articles varied between programmes. Under the AFGP, FMSP and PHFRP programmes publication data was available between 1990-2004, and for the PSP programme publication data was available between 1995-2004, however, citation analysis for all 4 of these programmes was only carried out for articles published post and including 1997. For CPP, projects were only sampled between 2001-2004, according to information provided by the programme managers and the specialist and for FRP citation analysis was only carried out for 9 sample projects.

- 2.5.5 Each of the Specialist Reviewers was asked to make use of a numerical scoring system to summarise the different attributes of the programme for the benefit of the Core Team using a full scale of 1 to 10. These figures are internally consistent but cannot be used to provide a global summary table giving comparison across programmes as the scoring system could not be moderated in the time available. It was therefore agreed that although these numerical scores would reflect each Specialist's report, they would not be amalgamated. The limitations of the numerical information were made explicit to Programme Managers and others to avoid protracted debate on the "scores" given.
- 2.5.6 Programmes were scored 1 (none/worst) – 10 (much/best) for the following:
1. To what extent the programme contributes to new knowledge;
  2. To what extent the programme uses existing knowledge creatively in new contexts;
  3. 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;
  4. Demonstrate awareness of all current knowledge (journals, books, web-based information) including in developing country literature, English language literature and non-English language literature;
  5. Extent to which the expected science achievements outlined in the log frame been met (key projects, outputs at programme level);
  6. Extent to which projects and the programme have contributed to science capacity building in the scientific communities in developing countries;
  7. Development of long-term institutional relationships between UK institutions and Southern institutions;
  8. Rating of the overall result knowledge dissemination from programme:
    - a) To science community (refereed, non-refereed, web-based, other media),
    - b) To developing country policy audiences;
    - c) To developing country outreach services;
    - d) To developing country end users (farmers, foresters, fisher folk.);
    - e) To the international donor community.

## **2.6 Criteria for Evaluation of Impact on Livelihoods**

- 2.6.1 The impact of RNRRS programmes is examined from a number of perspectives:
- The direct impact that individual projects, or clusters of projects, have had on poor people, either through increasing the opportunities open to them, by providing new technologies, reducing the risks they face, through reduction of pest and disease, or changing the policy environment within which they have to operate.
  - The impact made through influencing partners; partners in the international research community, in the national research system, in the national public service provision sector, in the private sector and in the donor community
  - The potential impact that projects and programmes might have in the future.
- 2.6.2 There were considerable challenges in undertaking this assessment given the resource constraints and tight time frame. It was particularly difficult to get any feel for impact of the more recent projects and clusters as many analysts would argue it is difficult to assess impact on poverty in less than ten to fifteen years after the start of a project.
- 2.6.3 From the beginning, it was recognised that the evaluation would not be able to conduct any new assessment exercises, but would be reliant on secondary studies, data on project performance and information on project and programme processes. Because of the number of projects undertaken in the RNRRS, it was also accepted that the emphasis would be on programme level information, rather than project level.

2.6.4 Where existing evaluation studies allowed, poverty was analysed, as much as possible, using the Sustainable Livelihoods (SL) Framework adopted by the Rural Livelihoods Department (RLD) of DFID in the late 1990s. This is in contrast to assessments of impact of agricultural research that focus on changes in productivity of crops and livestock. An SL approach implies looking at changes in economic poverty, equity issues, gender impact, environmental sustainability, vulnerability, social capital and empowerment, capacity building and policy impact.

## **2.7 Criteria for evaluation of Programme Governance and Management**

2.7.1 The TOR for the evaluation identify explicit governance and management issues in relation to project cycle management and programme management, as well as more general issues, particularly related to capacity building and skills transfer, and influencing activities.

2.7.2 Although the TOR do not make explicit reference to some important areas, the Core Team clarified with DFID that the following should also be considered:

- Expected and actual roles and responsibilities of PMs, and how these vary across the programmes.
- Strategies and plans prepared by PMs beyond the programme log frames – comparing and contrasting approaches, structures, balance and strategic planning processes.
- Management systems and processes adopted by programmes, especially for participation, executive decision-making, communications, reporting, monitoring and review, and human and financial resources management.
- Programme governance, particularly the effectiveness of the PAC arrangements with respect to strategic guidance, transparency and accountability.
- DFID’s strategic governance and management of the RNRRS as a whole, including arrangements for strategic direction, monitoring and evaluation, and utilising the outputs of the strategy.

2.7.3 The governance and management evaluation criteria used for the evaluation are concerned with “fit for purpose”. Essentially, the Core Team evaluated the extent to which the governance and management structures and processes are appropriate, and whether they are actually delivering programme contracts economically, efficiently and effectively. They examined the planned and actual management roles and processes operated by DFID and programme managers, identifying points of good practice and issues requiring further consideration in designing the successor to the RNRRS.

## **2.8 Challenges faced by the Evaluation**

2.8.1 The evaluation team faced a number of challenges which provide lessons for future evaluation activities.

- Although there was sufficient total time available for the evaluation, the overall deadline meant that there was insufficient time between the start and finish dates to allow for reflection, repeat consultations, and refinement of conclusions through dialogue.
- Even given the time available, it was never going to be possible to do more than sample the more than 2,000 projects and other studies funded and the copious documentation relating to each of these. This made the sampling system of critical importance to securing a comprehensive overview of each programme.
- A major limitation was the lack of independent or objective performance data on which the evaluation team could draw. This applied at every level from monitoring of project performance, evaluation and impact assessment, through to assessment of programme achievements by programme managers and by DFID. Even the impact assessment activities carried out by PARC relied on self-assessment by programme managers.

- Despite all of the efforts to collate and make available information held at the centre or by programme managers, the available data still displayed a number of limitations, in terms of volume, completeness, accuracy, and currency. Specific examples are: lack of comparability between programmes' annual reports despite the guidelines on contents and formats; out of date contact and address lists; absence of a reliable central index of documents indicating their location; and the variety of media types available.
  - The limited numbers of programme and strategy level documents meant that it was necessary to draw heavily on organisational memory to understand the evolution of the strategy, and changing priorities and processes. At the same time and in some cases the lack of continuity of personnel meant that this was an unreliable source.
- 2.8.2 To facilitate the evaluation of future programmes, careful consideration should be given to the collection and archiving of documentation, and to regular commissioning of objective reviews within an overall system of monitoring and impact assessments. There needs to be adequate investment in the evaluation process to support and manage change.
- 2.8.3 This evaluation is presented as a coherently argued interpretation of the DFID RNRRS Programme since 1994. It has, nevertheless, been constrained by the time available to the team and the sheer impracticality of detailed study of all the more than 2000 projects that have been funded and of the detailed documentation relating to each programme.