



## **How to evaluate the impact of environmental research on policy: CASE STUDIES**

**Output from SKEP Call 2 - Science to Policy Process: Lot 1**

Understanding the Impact of Environmental Research on Policy – Developing a  
Framework for Research Impact Evaluation and Guidelines for its Use



---

Policy Studies Institute



vrije Universiteit amsterdam



## *How to evaluate the impact of research on environmental policy: Case Studies*

Version 1.1, May 2010

### **Output from SKEP Call 2 - Science to Policy Process: Lot 1**

Understanding the Impact of Environmental Research on Policy – Developing a Framework for Research Impact Evaluation and Guidelines for its Use

### **Project Ref: 21205**

**This project was funded by the SKEP Network:** a partnership of government ministries, environmental protection agencies and research councils responsible for funding environmental research. It is an independent network, supported by member subscriptions. Further information on the Network is available from [www.skep-network.eu](http://www.skep-network.eu).

This report was written and edited by Ben Shaw and Sarah Bell of Policy Studies Institute, London based on the case studies conducted by the project consortium consisting of staff from the Institute for Environmental Studies (IVM) (The Netherlands), *Technologie voor Ecologie en Economie (VITO)* (Belgium) and researchers working in a personal capacity: Dr Annette Boaz, King's College London, Dr John Holmes, University of Oxford and Louise Shaxson, independent science policy consultant.

If you have any comments or questions on these case studies please send them to:

Ben Shaw  
Head of Environment Group  
Policy Studies Institute  
50 Hanson Street, London, W1W 6UP, United Kingdom  
Email: [b.shaw@psi.org.uk](mailto:b.shaw@psi.org.uk)  
Tel: +44 (0)20 7911 7544

# Contents

<b>1</b>	<b>Introduction</b> .....	<b>1</b>
1.1	Why were the case studies conducted? .....	1
1.2	What methods were used to identify and conduct the case studies?.....	1
1.3	Structure of report .....	2
<b>2</b>	<b>Case Studies - Summary Findings</b> .....	<b>4</b>
<b>3</b>	<b>Key findings from the case studies on evaluating research impact.</b> .....	<b>10</b>
3.1	Diversity in the objectives, scope and scale of research evaluation .....	10
3.2	Types of evaluation approaches employed .....	10
3.3	Generic challenges faced in evaluating the impact of research on policy .....	11
3.4	Challenges faced specifically with respect to the evaluation of the impact of environmental research on policy .....	13
3.5	Case study conclusions .....	14
<b>4</b>	<b>Full Case Studies</b> .....	<b>15</b>
	<b>Case Study 1: Stockholm Environment Institute</b> .....	<b>16</b>
	<b>Case Study 2: Irish Environmental Protection Agency</b> .....	<b>22</b>
	<b>Case Study 3: Consultative Group on International Agricultural Research</b> .....	<b>27</b>
	<b>Case Study 4: Defra/Environment Agency FCERM Programme</b> .....	<b>34</b>
	<b>Case Study 5: The Finnish Environment Institute (SYKE)</b> .....	<b>38</b>
	<b>Case Study 6: Foundation for Research, Science and Technology, New Zealand</b> .....	<b>42</b>
	<b>Case Study 7: Environment Agency (England and Wales) Post-Project Appraisals.</b> ..	<b>46</b>
	<b>Case Study 8: Swedish Environmental Protection Agency</b> .....	<b>52</b>
	<b>Case Study 9: Land and Water Australia</b> .....	<b>57</b>

# 1 Introduction

## 1.1 Why were the case studies conducted?

These case studies are an output from the SKEP Call 2 - Science to Policy Process: Lot 1 project, 'An assessment of current approaches towards the evaluation of the uptake and impact of research projects and programmes by environmental policy-makers'. The aim of the project was to develop an approach for the evaluation of the implementation and uptake of environmental research and guidelines for its use by the SKEP Network.

A literature review was conducted at the outset of the project<sup>1</sup> which aimed to synthesise and review existing knowledge and experience of the evaluation of research impact on policy and to use this as a building block for developing an applied model for environmental research evaluation impact on policy for use by SKEP partners.

One of the findings to emerge at an early stage of the literature review is that there are few good examples or sources that consider the detail of the practical approaches that have been used to conduct research impact evaluations. We therefore decided to conduct a series of case studies to both better understand for ourselves and illustrate clearly to potential users the approaches that could be used to evaluate research impacts. The case studies have three aims:

1. To illustrate relatively concisely some of the approaches to evaluation that have been used to give non-experts a feeling for what is involved.
2. To draw lessons that can be applied to our recommendations for conceptual and practical approaches and guidelines on evaluating the impact of research on policy.
3. To contribute to a comparative review of the strengths and weaknesses of different research impact evaluation approaches.

The case studies have been published in this standalone report as supporting material for the two main outputs from this project which are *How to evaluate the impact of research on environmental policy: a seven step guide* and the *Guidelines and Supporting Information* that accompanies the Guide.

## 1.2 What methods were used to identify and conduct the case studies?

Two main approaches were used for identifying the case studies. The first involved the identification of potentially interesting studies mentioned in the literature. Often these were only alluded to or mentioned in passing and therefore required further desk-based research and/or interviews with key individuals involved in the evaluation of interest. The second

---

<sup>1</sup> The final Literature Review is contained in a separate accompanying report.

relied on recommendations from project partners and responses from the SKEP member survey conducted alongside the literature review. Of particular value was the work done by Dr John Holmes as part of SKEP Work Package 4 – ‘Dissemination and implementation of environmental research’<sup>2</sup> – which touched on existing research impact evaluation processes occurring in each of the SKEP member countries.

The selection of case studies chosen aimed to capture:

- Examples from a number of countries, both European and wider international examples.
- A variety of approaches for evaluating the impact of research on policy, ranging from those derived from theory to more practice-based approaches.
- A variety of scales, methods and resource intensities.
- Evaluations with different objectives, including those focusing on improvements in research management, to those intended at enhancing the delivery of impacts.

Once the case studies of interest had been identified, a standard set of questions were developed to be answered by each case study (see Table 1 on the following page) to ensure the material from each case study was comparable and in the same form. Any available written sources were initially consulted to answer as many of these questions as possible, and these were then clarified, expanded and checked via telephone interviews with appropriate individuals involved in each evaluation.

The nine case studies conducted are:

1. Stockholm Environment Institute,
2. Irish Environmental Protection Agency,
3. Consultative Group on International Agricultural Research,
4. Defra/Environment Agency for England and Wales (FCERM) Programme,
5. The Finnish Environment Institute (SYKE),
6. Foundation for Research, Science and Technology, New Zealand,
7. Environment Agency for England and Wales, Post-Project Appraisals,
8. Swedish Environmental Protection Agency,
9. Land and Water Australia.

### **1.3 Structure of report**

Summary details of the case studies are given in section 2. Section 3 details the findings to be derived from the case studies in relation to developing research impact evaluation approaches and Section 4 gives the full details of each the case studies.

---

<sup>2</sup> Holmes, J. and Savgard, J. (2008) ‘Dissemination and implementation of environmental research’ *SKEP Work Package 4* [online]. Available from: [www.skep-era.net/site/files/WP4\\_final%20report.pdf](http://www.skep-era.net/site/files/WP4_final%20report.pdf)

**Table 1: Key questions informing the case study analysis**

Question	<i>Examples of follow up questions</i>
1. What is the background to the development and use of evaluation in your organisation?	<ul style="list-style-type: none"> <li>- Why are you interested in research evaluation?</li> <li>- How did the need for evaluation develop?</li> <li>- Did any specific papers/models help to inform/shape your approach to evaluation?</li> </ul>
2. What was the objective of the research evaluation?	E.g. Quality assessment? Learning about research management or research policy interaction? Assessment of value for money?
3. What impacts are you interested in evaluating?	<ul style="list-style-type: none"> <li>- Type of impact</li> <li>- Duration of impact</li> <li>- Strength of impact</li> </ul>
4. What types of methods were used to assess the policy or regulatory impact of research?	E.g. interviews with policy users or researchers, surveys, citation searches etc?
5. What, if any, other evaluation approaches did you consider using before developing this framework?	Were there any reasons for not using these and developing a new framework?
6. How would you change/develop the framework if you were to apply it again?	Is this approach something that is still under development and in use? If so, is it evolving? How?
7. What advice would you give to somebody else considering the development of an evaluation framework, particularly one to assess the impact of research on policy uptake and implementation?	What lessons did you draw from developing/using your framework?
8. What were the benefits and drawbacks of the framework?	(if not already alluded to in previous questions)
9. What range of potential users did you need to understand and accommodate in the framework?	(explain that we are interested in this as we need to develop some practical guidelines for evaluation of relevance to users with a wide range of expertise of evaluation)
10. What do you think are the generic challenges of developing a suitable research impact evaluation framework?	<ul style="list-style-type: none"> <li>- What difficulties did you encounter and how did you address them?</li> </ul>
11. What challenges are particularly pertinent for developing frameworks to evaluate the impact of <i>environmental</i> research?	(explain that there seems to be much literature on the evaluation of impacts of health research, but very little on environmental research)
12. What were the overall findings of the evaluation in terms of policy-impact?	<ul style="list-style-type: none"> <li>- Did the research have a significant policy impact? Did it highlight factors that can facilitate research having a greater policy impact?</li> </ul>

## 2 Case Studies - Summary Findings

In this section the key findings of each case study are presented. Full details of case study outputs can be found in Section 4 below.

**Case Study 1 - Who:** Stockholm Environment Institute (SEI)

**Where:** Multiple Countries (study included seven case studies)

**What:** Projects, programmes, processes of impact, knowledge support activities.

**Theoretical underpinning:** The approach used draws on four central ideas: types of policy impact based on Weiss' categories of policy impact (see full case study for details); depth of impact; the policy cycle and policy-research interactions; and factors known to enable or inhibit impact.

**Methods used:** In-depth key informant interviews were conducted with both supply-side and demand-side actors, covering a range of questions linked to the four central theoretical ideas.

**Methodological issues encountered:** Attribution; subjectivity of respondents' recall; timescale for impacts to occur.

**Case Study 2 - Who:** Irish Environmental Protection Agency (IEPA)

**Where:** Ireland

**What:** Programmes

**Theoretical underpinning:** Evaluation frameworks used by a number of organisations were explored, but the framework adopted is most closely aligned to the Australian Research Quality Framework. It focuses on both research quality (input, output and reputational quality) and research impact (policy, commercial and collaborative impact).

**Methods used:** Collation and assessment of data on sub-criteria covering research impact and quality. Data from interviews, surveys, citation analysis, collation of proxy measures, and documentary analysis. Sub-criteria assessed by expert panel to give ratings of research quality and impact.

**Methodological issues encountered:** Documentation required for evidence of impact can be poor; resource intensive approach; mistrust of researchers of the motivations behind the

evaluation; recall bias; timing; uncertainties associated with environmental research can affect policy-maker willingness to act on research findings.

**Case Study 3 - Who:** Consultative Group on International Agricultural Research (CGIAR)

**Where:** Multiple Countries

**What:** Projects, programmes

**Theoretical underpinning:** The approach draws on various ideas including: types of policy impact (based on work by Weiss as mentioned in SEI case study); hypothesised impact pathways; resource allocation; theories of influence.

**Methods used:** Creation of hypothetical impact pathways; testing with key informants and iterative approach to interviews to address subjectivity; comparison with a counterfactual; quantification of impacts and calculation of economic indices of impact and return on investment.

**Methodological issues encountered:** Reliability of key-informant interviews; poor documentation of activities undertaken to enhance impact; balancing breadth and depth of interviews; moving beyond traditional impact measures (e.g. citation analyses that do not correlate well with policy impact) and finding indicators able to measure less tangible impacts; the Cassandra problem (see full case study) timing and duration of impacts; attribution.

**Case Study 4 - Who:** Defra and the Environment Agency for England and Wales

**Where:** England and Wales

**What:** Projects and themes in broader programme.

**Theoretical underpinning:** Practically derived rather than underpinned by theory. It is a system-based approach which embeds monitoring of objectives, outputs and outcomes into project and programme management, from initiation to post-project, with proportionate degrees of effort.

**Methods used:** Central component of evaluation relies on assessing progress towards goals a SMART Benefits Delivery Plan developed by the project officer at project inception. A Programme Benefits Realisation Plan is also maintained - including a complete description of benefit, time-scales and evidence of benefits realisation – which is reviewed annually. Key themes explored include: the development of the knowledge base; policy, process and operational outcomes; and cost savings and resources.

**Methodological issues encountered:** Need for a common reporting and assessment system to be used on an annual cycle as a matter of routine.

**Case Study 5 - Who:** The Finnish Environment Institute (SYKE)

**Where:** Finland

**What:** Programmes, projects

**Theoretical underpinning:** Primarily a practice-based approach, but two ideas have shaped the methods used; triangulation of methodologies and multi-layer evaluations (i.e. different timescales – 10, 4 year and annual evaluations), and intervention theory (considering whether the research-based justification for the policy intervention at the outset proved valid)

**Methods used:** Ten-year evaluation - conducted by an independent evaluation panel with inputs of data from interviews with over 200 potential users of SYKE's work. Four-year evaluation - comprised of internal reviews that are cross-checked between programmes and with external stakeholder interviews and surveys. Annual evaluations - involve reporting on impact and effectiveness and a balanced scorecard approach. Staff self-evaluations of individual performance against organisational goals.

**Methodological issues encountered:** Involving potential users from the outset of the research design process; maintaining awareness of the political context to science and longer-term science needs; attribution; timing; uncertainties associated with environmental research can affect willingness to act on research findings.

**Case Study 6 - Who:** The Foundation for Research, Science and Technology (FRST)

**Where:** New Zealand

**What:** Thematic research programmes

**Theoretical underpinning:** Based primarily on practical considerations and advice from evaluation consultants rather than conceptual/theoretical models.

**Methods used:** Three methods: key informant interviews with research users to provide depth of information; a web-based questionnaire to provide breadth; and programme case studies which are researcher-focused but with external stakeholder input. Questions asked in the interviews and surveys aim to determine level of awareness of the research amongst potential users, relationships between FRST researchers and users, nature of benefits obtained from the research, rating of usefulness of FRST research for achieving benefit, rating of impact of each benefit on user organisations, and barriers/gaps in the research.

**Methodological issues encountered:** Attribution; diversity of users and research such that not all research will be perceived to be of value; mechanisms for improving the links

between researchers and users to improve uptake; diversity of users and impact pathways for different types of environmental research.

**Case Study 7 - Who:** Environment Agency for England and Wales

**Where:** England and Wales

**What:** Projects

**Theoretical underpinning:** Practically derived.

**Methods used:** Workshop format with key players and Agency project managers, covering: project background, overview benefits delivery, project processes, and conclusions. Only a sample of the existing projects were evaluated – approx 5-10 out of 300.

**Methodological issues encountered:** Need for clarity on outputs and outcomes; specification of intended benefits at the outset; responsibility for post-project monitoring; staff turnover; staff buy-in to the process; timing; need for evaluation to reflect the interdisciplinary nature of environmental research.

**Case Study 8 - Who:** Swedish Environmental Protection Agency (SwEPA)

**Where:** Sweden

**What:** Projects (ad-hoc), and quality, relevance and management of research

**Theoretical underpinning:** None cited – evaluation of the impact of research on policy tends to happen in an ad-hoc way, or is considered in planning cycles when considering why a particular environmental goal has or has not been achieved.

**Methods used:** Identification of research that has contributed to national environmental goals; bibliometric analyses two years after project completion. Ad-hoc evaluations are used to gain insights using a variety of methods e.g. interviews, tracking direct use of research in policy negotiations.

**Methodological issues encountered:** Need for standard methods to allow comparison of different projects and programmes – these could be supplemented with more ad-hoc evaluations tailored to needs and circumstances; need to involve users at an earlier stage; poor availability of data.

**Case Study 9 - Who:** Land and Water Australia

**Where:** Australia

**What:** Projects

**Theoretical underpinning:** Key aspects of the approach are the management of knowledge assets post-project completion (no projects are considered closed), and the integration of user needs into research programme design, management, communication, monitoring and evaluation.

**Methods used:** The approach aims to understand research impacts *and* policy-makers' knowledge-seeking behaviour. LWA evaluation guidelines suggest a range of possible methods – e.g. targeted stakeholder surveys, narratives, records kept by research teams, cost-benefit analyses - the choice of which will depend on the evaluation questions chosen.

**Methodological issues encountered:** Attribution issues, timing of evaluation, ensuring evaluation is integral to programmes rather than an isolated or retrospective process.

Whilst Land and Water Australia has recently been abolished, their proposed approach was interesting as it explicitly goes beyond assessing research impact, combining it with an assessment of how to *enhance* impact.

### **3 Key findings from the case studies on evaluating research impact.**

A number of important findings can be drawn from this case study analysis, primarily concerning: the considerable diversity in the objectives, scope and scale of research evaluations conducted; the types of methods employed; the challenges faced; and the challenges faced specifically with respect to the evaluation of the impact of *environmental* research on policy. This section explores each of these in more depth.

#### **3.1 Diversity in the objectives, scope and scale of research evaluation**

Most of the impact evaluations conducted had multiple objectives, including accountability to funding organisations and justification of their investment, and improving the effectiveness, relevance and wider impacts of the research programme (learning). For some, such as the CGIAR, the accountability side was paramount and therefore economic valuation of impact was a high priority. For others, learning was more important. The SEI, for example, was more interested in unpacking the concept of policy impact; what it is, and how can it be measured and enhanced.

Most of the case studies explored programme-level evaluations, though the SEI and CGIAR focused on a number of individual projects. The scope and scale of the research impact evaluation largely depended on the resources available, ranging from an ad-hoc non-structured approach in the Swedish Environmental Protection Agency, to the incorporation of research impact as a consideration in wider evaluation processes as seen at SYKE and the Irish EPA, to much more in-depth focused approaches at the CGIAR and SEI. Whilst some organisations focus primarily on research uptake and management processes (SwEPA and EA), others attempt to explore wider impacts and more long-term outcomes (New Zealand FRST, SEI, CGIAR, SYKE).

#### **3.2 Types of evaluation approaches employed**

The approaches used for evaluating the impact of research on policy were most commonly derived from practical need and tended to be process-driven, with the exception of the more academic organisations such as the SEI and CGIAR, who drew more heavily on existing theoretical and conceptual frameworks.

Most of the case studies employed the use of key informant interviews with researchers and/or intended policy users. In some cases these were accompanied by surveys to broaden the reach of the evaluation, as in the New Zealand case study, or by bibliometric and documentary analyses to supplement the information gained from the interviews, or to help identify the key informants. The Environment Agency for England and Wales tested a slightly less resource-intensive approach, choosing to hold workshops with key players and

Agency project managers to consider and discuss the benefits delivery of the research projects in question. In general it was felt that multiple methods were needed to overcome many of the challenges associated with evaluating the impact of research on policy, particularly for environmental research.

### **3.3 Generic challenges faced in evaluating the impact of research on policy**

These case studies highlighted numerous challenges faced when evaluating the impact of research on policy, and raised a number of mechanisms to try to address such challenges.

First, and most commonly raised, is the issue of **attribution**; how to measure the impacts specifically of research on policy when so many factors influence the policy process, and furthermore, how to attribute impacts to one particular research project or research organisation when policy decisions tend to be based on the culmination of findings from various projects and programmes. CGIAR suggests this issue may be addressed in three ways:

- The use of a counterfactual to identify what the policy outcome would have been without the input of a specific piece of research (though developing a counterfactual can itself prove challenging).
- Demand-side approaches to impact evaluation (as opposed to supply-side), which use major policy events as the starting point and work retrospectively to establish the separate influences of the many research suppliers and other factors on the policy responses.
- Institutionalising the process of impact evaluation, i.e. ensuring that staff and management take responsibility for recording outputs, outcomes/influences and policy responses related to their research (as illustrated by IFPRI in the CGIAR case study). Independent evaluators can verify these and translate them into meaningful measures of their impact on economic welfare, assessing to what extent any changes in welfare can be attributed to policy research institutions and their partners.

Second is the issue of the **appropriate timing** for the evaluation process and how to capture the duration of research impact. If the evaluation is conducted too soon after the completion of the research, no impacts may yet be apparent. However, research results can set the stage for a policy change that occurs at a much later date when the original research results are forgotten, i.e. if the evaluation is conducted too late, the link between the earlier research and the substantially later policy change is broken, and key informants may no longer make the connection. The Environment Agency for England and Wales therefore suggested a two-stage process for future evaluations; the first stage would rely on documentation and a small number of interviews shortly after project completion, and the

second stage would adopt the workshop format and would be held once potential benefits of the project are expected to have emerged.

Related to this is the third issue, the **reliability of information from key informant interviews**; some key informants will have been much more closely engaged with the research than others, some may play more central roles in the development of policy, some may have more or less reliable and objective recall of key conversations and critical events, and some will have changed posts and lost contact with the programme/policy in question before final decisions were made. The New Zealand case study used the interesting approach of using independent consultants to help identify appropriate key informants. They focused on selecting 'key players' or 'information nodes' rather than drawing from the general pool identified by researchers. The consultants used a matrix template with the type of research on one axis and type of user organisations on the other axis to ensure a comprehensive coverage of users.

Fourthly is the challenge in determining **which methods to use to capture as many impacts as possible**. There is a need to look beyond the use of traditional research impact indicators such as the number of publications in leading journals and subsequent citations, as studies (such as those by CGIAR) increasingly indicate that this measure does not correlate well with policy influence or wider impact. Indicators need to be developed that can identify less tangible impacts, such as changes in common ways of examining policy processes that lead to multiple changes in policy decisions across countries, institutions, and/or individuals.

These sorts of impacts are difficult to trace and capture, and doing so becomes increasingly difficult beyond the project level as more actors become involved, together with exponentially greater information sources and motivations for use. SYKE highlighted the particular difficulties in tracing uptake and impact of 'blue skies' research, as research planned and executed without the input of or initial demand from policy-makers is rarely taken up directly or immediately by policy-makers.

Linked to this is the '**Cassandra problem**' (Ryan and Garrett, 2004); is it possible to capture and value 'good research advice' that is not taken? Or delays in taking the advice? It is suggested that an evaluator could calculate the opportunity costs of a 'wrong decision' (i.e. estimate the cost of the alternative to not taking the advice). Alternatively, so-called 'good advice' may have arisen from flawed research, in which case it is to the policy-makers' benefit if they did not choose to accept it.

Finally there is the challenge of the **resource intensive nature of impact evaluations**. Due to the many challenges associated with such evaluations, evaluators often try to employ multiple methods in order to identify the impacts from several perspectives. There is a need, however, to find the best approach based on the limited resources available.

The New Zealand case study, for example, tries to find the best balance between conducting in-depth interviews with a few key informants and facilitating wider engagement through the use of written survey questionnaires (the latter is less resource intensive and enables contact with more research users, but also potentially provides less details and credibility).

The Irish EPA highlighted that many organisations have conducted impact evaluations but many fail to document these processes in detail, thereby forcing the evaluator to move beyond straightforward documentary analysis. They also found that the scientific and academic community were often not particularly interested in answering questions about the wider impacts of their research, such that considerable time and effort was required to meet the researchers involved on an individual basis to explore and highlight the benefits of evaluating the impact of research and thereby encourage cooperation. The Environment Agency for England and Wales evaluation found that it was the senior staff on the customer side that were often more reluctant to participate in the evaluation process, although where they did attend the workshops, they found them to be enjoyable and beneficial.

### **3.4 Challenges faced specifically with respect to the evaluation of the impact of environmental research on policy**

A number of research impact evaluation challenges are particularly pertinent to environmental research.

Environmental research programmes are often very **diverse** and may cover wide-ranging subject matter – e.g. from research targeted at the land environment, to the marine, to more people-based environmental behaviours research. The potential user population may therefore be very diverse, such that not all the research in a programme will be of value to everyone. The impact pathways may be also very different for different types of environmental research, which can complicate efforts to develop generic impact evaluation criteria suitable for all projects funded by a particular programme that will still capture the level of detail required in each evaluation. A distinctive characteristic and further complicating factor linked to the evaluation of environmental research is that it is often interdisciplinary, and increasingly brings together the natural and social sciences. The evaluation approach needs therefore to reflect this interdisciplinarity.

A further challenge is that environmental research is likely to produce **environmental benefits for society**. It can be difficult to quantify such benefits as they tend to lack market values, and the development of effective environmental valuation techniques continues to prove difficult. Similarly, **intangible outcomes** produced by much environmental research are considered to be very valuable but these are also more difficult to measure. In the New Zealand case study, for example, users regarded new knowledge which might lead to a change in awareness or understanding of an environmental problem as being important to

the development of a new management tool. This is quite different to many other research areas, where the main outcomes are tangible benefits (e.g. manufacturing research). They addressed this by ensuring questions were included in the web surveys and key informant interviews to capture these benefits and their impact, and by following these up with a cost benefit study.

SEI suggested that a particular challenge in assessing the impact of environmental research is that the findings of much environmental research **will only have a significant ultimate impact if policy-makers act on them *immediately***, which rarely happens. Due to complex feedback systems and lag times associated with climate change science, for example, research into the effectiveness of climate change mitigation measures may recommend mitigation measures that will ensure climate change targets are met, but only if implemented immediately. It is therefore important to assess the actual impact of environmental research rather than its *potential* impact.

A significant reason why policy-makers do not tend to act on environmental research recommendations immediately relates to the **considerable uncertainties and caveats** often associated with environmental research (e.g. climate change research). Impact evaluations may therefore need to consider how these uncertainties have been explained to the policy community, to fully understand how uncertainties may have compromised the potential uptake of the research. In addition, policy-makers need to be honest in how they communicate uncertainties to ensure success of the policy in question.

### **3.5 Case study conclusions**

These case study analyses highlight the diversity of existing approaches to the evaluation of the impact of research on policy, ranging from relatively basic largely practically-derived approaches, to more in-depth frameworks grounded by conceptual and theoretical insights. The approach used can be tailored to the needs and available resources of the organisation in question, but must be closely tied to the objectives of the evaluation and must attempt to address the key challenges highlighted in this section.

## 4 Full Case Studies

Details of the nine case studies conducted are given below under six standard headings for each case study:

1. Name and background of organisation.
2. What is the background to this study?
3. Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?
4. What evaluation approach was used to assess policy impact?
5. Were any challenges highlighted in assessing the impact of the organisation's work (and any particularly relating to the impact of *environmental* research)?
6. Information sources for case study.

## Case Study 1: Stockholm Environment Institute

**Name and background of organisation:** Stockholm Environment Institute (SEI).

The Stockholm Environment Institute (SEI) is an independent, international research institute conducting research into sustainable development and environment issues. It works at local, national, regional and global policy levels, and aims to clarify the requirements, strategies and policies for a transition to sustainability.

SEI's mission is '*to support decision-making and induce change towards sustainable development around the world by providing integrative knowledge that bridges science and policy in the field of environment and development*' [SEI website, 2009].

### What is the background to this study?

The SEI recognises that as an organisation with an important policy audience at global, national and sub-national levels, their work should primarily be evaluated on the basis of what types of public policy impacts it generates at these different levels. This study was therefore initiated to unpack the concept of policy impact and to consider a number of questions about its evaluation:

- How and where might one observe policy impact?
- How might that impact be traced back to the introduction of sustainability knowledge to influence actors' orientations in new, more sustainable ways?
- What different kinds and different levels of policy impact can there be?
- What types of factors around the process facilitate or impede policy impact?

The insights of the study are built upon an analysis of case studies where SEI has provided knowledge support to policy formulation or implementation. The aim of these case studies is to identify and characterise the policy impact and to examine the contributing factors and external conditions under which the policy advice and support was provided.

The case studies included:

1. *Policy analysis for state-level climate action in the US*: examines SEI's involvement in the 2007 Washington State Climate Advisory Team (CAT) process.
2. *Formulating policy advice on reducing the carbon footprint of housing*: the Future Sustainability group at SEI developed a scenario tool – REAP (Resource and Energy Analysis Programme) – with which SEI creates scenarios that can be used to help compare the ecological (or carbon) footprint impact of various policy interventions. This case study looks at the application of the tool in the environmental assessment of the Leeds city region housing policy within the Yorkshire and Humber Regional Spatial Strategy and the impact and resulting spin-offs from that piece of work.

3. *Waste policy in Estonia*: examines SEI's influence on the process of developing the legal basis and economic instruments for waste management in Estonia between 2000 and 2008, and analyses SEI's multi-faceted role in developing and implementing the legislation regarding packaging and packaging waste, in developing the National Waste Plan, and in contributing to the proposal for waste taxation.
4. *Introduction of environmental tax reform (ETR) in Estonia*: explores the role of the SEI's Tallinn Centre in policy formulation and stakeholder engagement, from the start of the ETR process.
5. *Policy impact of Sustainability Knowledge – the Ozone Layer Protection Programme*: evaluates the impact of the knowledge and support provided by SEI to inform the Swedish government's Ozone Layer Protection programme (OLP), a programme designed to provide direct bilateral assistance to developing countries for the 'sustainable and cost-efficient phase-out of ozone depleting substances'.
6. *Tropospheric ozone case study*: this case study both examines the integration of specific SEI-generated scientific knowledge into the LRTAP (Long Range Transboundary Air Pollution) policy context, and explores the impact of policy advice provided by the SEI to inform the development of policy instruments i.e. it explores how the overall capacity and legitimacy of SEI's Atmospheric Environment Group have allowed its studies to be taken into account in policy arenas.

Each case-study write-up provides a background and introduction to the case, followed by a brief discussion of the policy context and the specific policy issues where SEI and its knowledge co-creation partners have actively provided knowledge and support. Finally, for each case study there is a discussion about the conditions that determined the output (and in some cases outcome) of the policy support.

**Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

Four central ideas informed the evaluation approach used in this study:

1. *Types of Policy Impact*. The work of Carol Weiss (1979)<sup>3</sup> provided a useful starting point for the exploration of the concept of 'policy impact'. Weiss identifies seven different ways in which research may be used in policy and emphasises that research utilisation does not simply mean the citing of a particular piece of research in a policy document or the 'problem-solving' approach, but also has other less tangible impacts e.g. enlightenment, interactive, political, tactical and research as part of the intellectual enterprise of the Society.

---

<sup>3</sup> Weiss, Carol H. 1979. *The Many Meanings of Research Utilization*, *Public Administration Review*, 39 (5): 426-431.

2. *Depth of Impact*. This idea is based on the recognition that a sequence of learning elements can provide a means for categorising the depth of impact in terms of learning (Nilsson 2006 adapted from Huber, 1991) e.g.
  - *Knowledge acquired*: an assimilation of experience, as well as new ideas and concepts from other actors.
  - *Knowledge interpreted*: gaining new understandings of cause-effect relations of policy problems and how to resolve them.
  - *Knowledge institutionalised*: incorporation of new knowledge into procedures, rules, policies and other tangible outputs for implementation.
3. *The Policy Cycle*. Research institutes such as the SEI will have most impact through ensuring timely delivery/communication of outputs and through feeding into multiple stages of the policy cycle. Whilst a simplified policy-cycle is considered to have four broad stages – formulation, translation, implementation, evaluation – in reality it is more complicated. Strategic policy tends to be formulated at a relatively high level of governance (regional or international e.g. EU, UN), which is then translated into policy measures at the national or state level, and finally implemented at the local level. When trying to increase policy impact, it is therefore useful to remember that important policy decisions are made and can be influenced by the input of sustainability knowledge at each of these levels – regional strategic, national measures and local implementation schemes.
4. *Factors that influence potential for impact* In addition to these three ideas, it was recognised that the extent of policy impact of research would also depend on a number of additional qualities, including:
  - Quality of the scientific/technical work.
  - Effectiveness of user engagement processes at the science-policy interface
  - Communication efforts
  - External context

### **What evaluation approach was used to assess policy impact?**

The policy impacts of each of the SEI workstreams explored by the case studies are largely evaluated through the use of semi-structured key informant interviews with both supply-side and demand-side actors (with a particular focus on the latter). In each case, a minimum of two supply-side and three demand-side actors were sought for interview. From the outset of each interview it was made clear that the aim of the interview was to improve the services provided by the SEI and the modes in which they supply information i.e. that the interviewers were looking for honest feedback about successes and failures in terms of knowledge impact, and for supporting evidence where successes were claimed.

The four ideas above – types of policy impact, depth of impact, the policy cycle, and factors influencing the potential for impact - informed the development of the following interview framework:



**Table 2. Interview framework**

Main question	Sub-question
<b>The Dependent Variable</b> (Questions about the learning or other impact that occurred)	
Where is the policy impact?	In what stage of the policy process? By what actors? (decision maker, bureaucrat, etc)
What kind of policy impact is this?	Can we categorise impact into: <ul style="list-style-type: none"> <li>- <i>Instrumental</i> – we can detect changed decisions</li> <li>- <i>Enlightenment</i> – we can detect new awareness and learning, and new ways of seeing and doing things</li> <li>- <i>Interactive</i> – we can detect increased interaction and engagement between actors</li> <li>- <i>Tactical</i> – we can detect knowledge used to delay or deflect attention</li> <li>- <i>Political</i> – we can detect an actor positioning, underpinning or undermining new knowledge based on pre-conceived notions</li> </ul>
How deep is this policy impact?	Can we categorise depth of impact into: <ul style="list-style-type: none"> <li>- Knowledge acquired?</li> <li>- Knowledge interpreted?</li> <li>- Knowledge institutionalised?</li> </ul>
<b>The Independent Variables</b>	
<i>Questions about explanations for the observed patterns of impact in the process surrounding the knowledge provision, communication and uptake)</i>	
Actor capacities, orientations and incentives	Who or what is an actor? What is their role? What is SEI’s role as sender of Sustainability Knowledge? What is the users’ role as receiver?
Relevance of knowledge - to needs and wants of recipient	What information did they need/want? Why did they need it (what is its relevance?) How did they need it and did SEI provide it in the required format? (i.e. appropriateness of medium)
<b>Questions about the timing and the message</b>	
Timing	Was it timely? Is there a policy cycle/research cycle disjoint or was it right message at the right time?
Knowledge features	How do decision makers perceive ‘useful’ evidence? Was the message appropriate and was it delivered in the correct and useful manner? e.g. phone call/50-page report Precision?

	Quality? Format?
Organisational arrangements for knowledge transfer	
Collaboration arrangements in the provision of knowledge	
<b>Other contextual influences</b>	
Media attention	
Political interest	
Risks and perceptions of risk, uncertainties	
Visibility of the issue	
Public opinion	

**Were any challenges highlighted in assessing the impact of the organisation’s work (and any particularly relating to the impact of *environmental* research)?**

During the interview, two main challenges were highlighted:

- The attribution issue: as the SEI tends to work in partnership with other organisations, key informants sometimes struggled to identify how much of the overall impact of the work could be attributed to the SEI alone.
- Reliance on subjective recall of key informants; in some of the case studies, the actors interviewed were still involved in the process whilst in other they are retired and reflecting back on processes and their perceptions of impact. However, including studies with this longer timescale had the benefit that the study was able to reflect on actual outcomes/impact of the research rather than just potential predicted outcomes.

During the interview, it was suggested that the main challenge specific to assessing the impact of environmental research is that the findings of much environmental research will only have a significant ultimate impact if policy-makers act on them *immediately*, which rarely happens e.g. research into the effectiveness of climate change mitigation measures may suggest mitigation measures that will ensure climate change targets are met, but only if implemented immediately. It is therefore important to assess the actual impact of the environmental research, rather than its *potential* impact.

**Information source for case study**

Interview with Dr. John Forrester (York, SEI), together with reference to his report ‘Getting to policy impact: Lessons from 20 years of bridging science and policy with sustainability knowledge’ (available from:

[sei-international.org/mediamanager/documents/Publications/Policy-institutions/policy-impact-report-forrester-et-al.pdf](http://sei-international.org/mediamanager/documents/Publications/Policy-institutions/policy-impact-report-forrester-et-al.pdf)).

## Case Study 2: Irish Environmental Protection Agency

**Name and background of organisation:** Irish Environmental Protection Agency (EPA).

The Irish EPA aims to protect the environment for the benefit of the Irish population, through the regulation and policing of activities that might otherwise cause pollution. They seek to gather reliable information on environmental trends so that necessary actions are taken.

The specific mission of the Irish EPA is *‘to protect and improve the natural environmental for present and future generations, taking into account the environmental, social and economic principles of sustainable development’* (Irish EPA website).

The organisation is responsible for: licensing and control of large-scale waste and industrial activities to ensure that they do not endanger human health or harm the environment; national environmental policing; monitoring, analysing and reporting on the environment; regulating Ireland’s greenhouse gas emissions; environmental research and development; strategic environmental assessment; environmental planning, education and guidance; and, proactive waste management.

### **What is the background to the study?**

The Irish Environmental Protection Agency has used several different indicators in evaluating the research produced as part of its ERDTI Programme. They developed the evaluation framework discussed in this case study in order to provide stakeholders with information on the quality and impacts of its funded research (accountability) and a structure to facilitate learning and further development of the research programme (learning).

As a pilot study, the framework was applied to the programme of aquatic research, the area that received the greatest share of support between 2000 and 2006. The primary aim of the aquatic research was to develop new and innovative methodologies to be used in the characterisation and monitoring of datasets and models, and to thereby support the national implementation of the Water Framework Directive.

### **Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

A number of models and alternative conceptual frameworks informed the development of the evaluation approach used. These included:

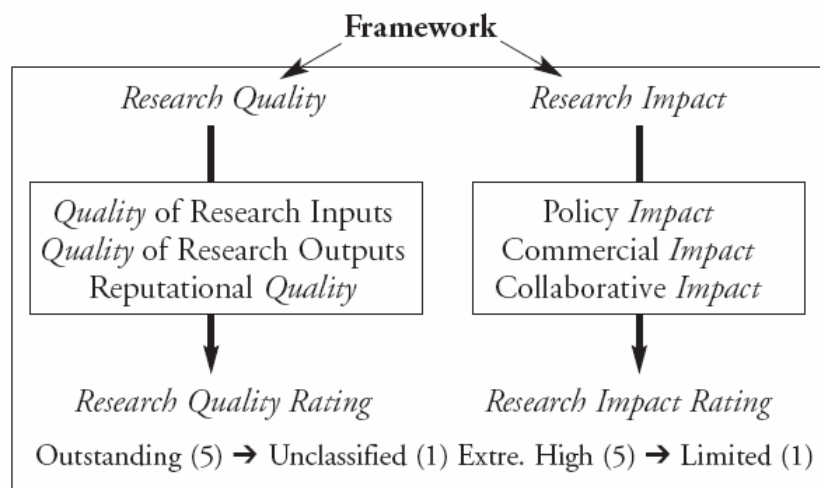
1. The Australian Research Quality Framework (2005)
2. The UK Research Assessment Exercise (2003)
3. The ‘Evaluation of the Environmental Output Class’ conducted by New Zealand’s Ministry of Research, Science and technology (2004)
4. Evaluation methods used by the Science Foundation Ireland (2001-2005)

5. Evaluation methods used by the Institute of Public Administration
6. Evaluation approaches used by the EU LIFE Programme
7. Evaluation approaches used by the Irish Health Research Board
8. Academic evaluation research conducted in the Netherlands.

**What evaluation approach was used to assess policy impact?**

After undertaking an extensive literature review and a number of in-depth interviews with experts working in the field of impact evaluation, the following framework was developed, with two main criteria; research quality and research impact. Each criterion is divided into three sub-criteria and a rating scale is then used to capture the assessment in a simple numerical format.

**Figure 1. The Framework**



*Ref: O’Leary, G., Boyle, R., Donlon, B. and Sheils, L. (2008): p162*

It appears that the overall assessment is based on data collected for the sub-criteria such as for research quality: the number of Masters/Doctorate students and the completion time for students; the number of publications, citations etc; and the contribution to work of national and international academic committees and working groups, etc. For research impact measures were used such as: work on advisory panels to industry or government agencies; the effects of the research on commercial activities such as patent applications or savings in resource use; and the linkages between the researchers and target stakeholders. The data required to assess the research against each of the sub-criteria is collected using a range of methods such as documentary analysis, citation analysis, interviews with project personnel and/or end-of-project questionnaires sent to project co-ordinators. A context statement is also completed for each project to allow for factors beyond these measures to be taken into account. This aspect builds in some element of flexibility to the fairly rigid framework.

For this case study, we are particularly interested in how ‘Policy Impact’ was measured. The primary measures for policy impact include:

- How the programme contributes to national and international policy development e.g. to the implementation of the Water Framework Directive, the Nitrates Directive, the Habitats Directive and Strategic Environmental Assessment.
- Whether the Government has applied programme findings to policy development
- Whether the programme has contributed to work on advisory panels to industry of Government agencies.

‘Collaborative impact’ is also an important criterion to consider as successful collaborative impacts can often lead to successful policy and commercial impacts.

The rating scale provides a summary of the assessment carried out of research against the agreed criteria and indicators. A five point rating scale (extremely high, very good, high, good, limited) is proposed for research impact, and the ratings were carried out by a team of three national experts (two in house and one external) with experience in Water Framework Directive compliance and Water Research (including the interviewee, Lisa Sheils).

In this pilot evaluation, it was found that 89 per cent of projects demonstrated some policy impact; that is, 62 per cent demonstrated a high level, whilst 17 and 10 per cent showed moderate and low impact respectively. The high-level of impact reflects the targeted effort towards the implementation of the Water Framework Directive.

This pilot evaluation formed the basis of a benchmarking and evaluation exercise which has now been mainstreamed into the EPA’s everyday project management processes such that progress on certain impact indicators are captured on a regular basis.

The framework informed the development of a tool that is constantly being applied and that is evolving as new needs and information become available. Subsequent use of the approach in the Environmental Technologies Research Programme, for example, has included the production of a series of case studies alongside the final impact and quality scores. These case studies were based on the responses to surveys sent to the principal investigators of 18 projects within the programme, asking about a number of payback categories, including: knowledge production, capacity development, informing policy and environmental benefits, and broad social and economic benefits (EPA, 2009).

**Were any challenges highlighted in assessing the impact of the organisation’s work (and any particularly relating to the impact of *environmental* research)?**

A number of general challenges were highlighted, including:

- Existing documentation of impact evaluation approaches: Impact evaluations are something which many organisations may have conducted but which they often fail to document in detail. Developing an informed conceptual framework for impact evaluation therefore requires the more labour-intensive approach of

speaking to relevant individuals in these organisations to identify the relevant information.

- Resource-intensity of the evaluation approach: The scientific and academic community were often not particularly interested in answering questions about the wider impacts of their research, such that considerable time and effort was required to meet the researchers involved on an individual basis to explore and highlight the benefits of evaluating the impact of research and thereby encourage co-operation. This approach also helped to overcome researcher mistrust about the motivations behind the impact evaluation; a number of researchers suspected that any evaluation being conducted to assess research impact and value for money may result in reductions in funding supplies.
- Evaluation criteria: Challenges were encountered in trying to develop sufficiently generic impact criteria that can be applied to multiple research programmes funded by the EPA whilst still capturing the level of detail required in each evaluation.
- Recall bias: Assessments of the impact of the research rely on the memory of the researchers and research-users associated with the aquatic research programme, many of whom have since moved on. This raises the issue of subjective recall/recall bias.
- Timing of evaluation: A balance must be found between ensuring enough time has passed after the completion of the research for impacts to have occurred, but not so much that the impacts can no longer be traced back to the original research. The ERDTI research programme co-ordinators are considering the re-evaluation of the policy impact sub-criteria after three years from this first evaluation in order to see whether policy impact has increased over time.

Environmental research programmes can be particularly challenging to evaluate as they often incorporate projects with very different types of expected outcomes; this can complicate efforts to develop generic impact evaluation criteria suitable for all projects funded by a particular programme.

#### **Information source(s) for the case study**

Interview with Dr. Lisa Sheils, in-house expert, Irish EPA.

O’Leary, G., Boyle, R., Donlon, B. and Sheils, L. (2008) ‘Providing a framework for accountability and learning in environmental research’ *Administration* 55(4): 159-170

Kivimaa, P. *et al.* (2008) ‘Approaches and practices in the evaluation of environmental research programmes’ *SKEP Work Package 3* [online]. Available from: [www.skep-era.net/site/files/WP3\\_Evaluation\\_Programmes.pdf](http://www.skep-era.net/site/files/WP3_Evaluation_Programmes.pdf)

EPA (2009) ‘Innovation for a green economy. Environment and Technology: A win-win story’ [online]. Available from: [www.epa.ie/downloads/pubs/research/tech/EPA%20STRIVE%20ET%20Report\\_Final.pdf](http://www.epa.ie/downloads/pubs/research/tech/EPA%20STRIVE%20ET%20Report_Final.pdf)

Wemaere, A. *et al.* (2009) 'An evaluation of the role of EPA research in the Water Framework Directive implementation in Ireland' *Biology and Environment: Proceedings of the Royal Irish Academy* 109B (3): 385 – 402.

## **Case Study 3: Consultative Group on International Agricultural Research**

**Name and background of organisation:** Consultative Group on International Agricultural Research (CGIAR).

CGIAR aims to ‘achieve sustainable food security and reduce poverty in developing countries through scientific research and research-related activities in the fields of agriculture, forestry, fisheries, policy and environment’ (*CGIAR website*).

CGIAR is a strategic partnership, whose 64 members support 15 international centres, working globally in collaboration with hundreds of government and civil society organisations as well as private businesses. One of CGIAR’s key priorities is to improve policies and facilitate institutional innovation.

### **What is the background to this study?**

This study aimed to build on a scoping study conducted in 2005/06 which assessed the extent and nature of existing documented impacts of policy-oriented research (POR) within the CGIAR. After the initial scoping study, the Science Council’s Standing Panel on Impact Assessment (SPIA) commissioned a number of additional case studies of the impact of POR, both to augment the evidence and to further the development of methodologies in the area of impact assessment.

Seven case studies were selected, with the aim of evaluating qualitatively *and* quantitatively the degree to which a policy change of interest could be attributed to a particular research project/programme or research-related output. The seven case studies chosen include:

1. The Impact of In-Trust Agreements on CGIAR Germplasm Exchange and the Role of Bioversity International in establishing the Agreements.
2. Assessing the impact of CIFOR’s influence on policy and practice in the Indonesian Pulp and Paper Sector.
3. Returns to POR: the case of barley fertilisation in Syria.
4. IFPRI and the Mexican PROGRESA Anti-Poverty and Human Resource Investment Conditional Cash Transfer Programme.
5. Policy change in dairy marketing in Kenya: Economic impact and pathways to influence from research.
6. Economic assessment of policy-oriented research on the private health costs of pesticide use in the Philippines.
7. Community-based fisheries management project in Bangladesh.

**Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

The study began with a review of existing approaches to the estimation and evaluation of the impact of policy-oriented research.

The components of an assessment of POR are itemised as the elements of an impact pathway, from resources committed to a research project or programme to the policy consequences of the research; inputs, outputs, dissemination (or uptake), influence and impact. This is also termed ‘theories of influence’.

The review draws on the work of Weiss (1979) to consider how research can be utilised by policy-makers (*uptake*), drawing particular attention to the Problem Solving, Political and Enlightenment Models. It also highlights research by Walter *et al.* (2003) and by Spilsbury and Kaimowitz (2002) indicating factors that may encourage research uptake.

The review describes three bodies of literature on approaches for measuring research *influence*:

1. Estimating the statistical relationships between output measures and influence measures.
2. Efforts to clarify the pathway through the policy process itself.
3. Research based on survey questionnaires/interviews with policy-makers themselves.

Finally, it considers existing approaches for measuring *impact*, including the use of Cost-Benefit Analysis for quantifying the difference made by a policy, the issues associated with trying to value non-market benefits of a particular policy change, and approaches for assessing the extent to which these benefits are attributable to the POR.

### **What evaluation approach was used to assess policy impact?**

#### *The CGIAR study*

While each case study is slightly different, the essence of the approach appears to be the construction of hypothesised impact pathways, in which the authors charted assumptions about where information from the research entered the policy-making system, the paths it took, and the decision-makers it influenced, and then sought to verify these assumptions of transmission and uptake. In most cases, this verification process involved conducting a large number of in-depth key informant interviews to explore influence and impact, together with the use of primary and secondary data sources and reports to validate and support the information gained from these interviews (the key informant approach can be useful where there are a limited number of people with significant enough depth and breadth of knowledge to speak informatively about what happened and why). Key informants were also used to broaden the scope of later interviews and add other interviewees to the original list. One study featured a round of feedback from the key informants so that they could comment on the perceptions of their colleagues to arrive at a more reliable subjective scoring of influence.

The second stage of the evaluation process involved quantifying the magnitude and distribution of the impacts of a particular policy change, and then generating a counterfactual to calculate what wider benefits of the particular policy would likely to have been generated with all the other players in the process still active, but with the research project in question removed. This helps to identify what proportion of the wider benefits of a particular policy may be attributed to the research.

The final stage of the evaluations was to compare the measured benefits attributed to the POR with the cost of the research to calculate the NPV, IRR and/or the BCR i.e. to identify the returns to the research investment.

#### *IFPRI's approach*

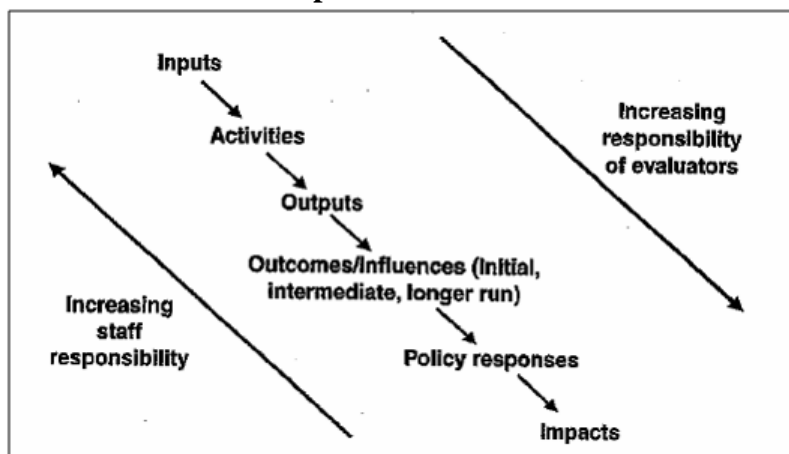
Ryan and Garrett (2004) specifically discuss IFPRI's approach to impact evaluation, which has evolved over the last 10 years. IFPRI, a research institute supported by CGIAR, undertook a number of case studies, beginning in 1998, adopting a mixture of quantitative and qualitative approaches to assess impact primarily at the project level, and primarily for accountability purposes. In 2000 it was decided that impact evaluations would be institutionalised and, in 2001, a number of pilot exercises were initiated involving ex ante impact evaluation on new projects. IFPRI also began to go beyond the project level to conduct evaluations of some of its thematic research programmes. By 2002-3, IFPRI decided to maximise the learning opportunities associated with impact evaluations and therefore to incorporate some aspects of impact evaluation in all its research activities. Through a series of focus groups, all research staff were requested to narrate instances where their research outputs had influenced policy and had subsequent social or economic impacts. Each focus group included staff from different research divisions in order to stimulate cross-fertilisation and validation.

By 2004, IFPRI was categorising products from economic policy research and related activities as:

- *Outputs*; activities or efforts that can be expressed quantitatively or qualitatively.
- *Outcomes/influences* (separated into initial, immediate and longer-term); measures of the use that clients or partners make of the outputs.
- *Policy responses*; imply a degree of attribution of the effects of the intermediate outputs and outcomes/influences on the formulation of new, or reinforcement of existing, policy.
- *Welfare impacts*; measurable effects of the attributed policy responses on the well-being of the ultimate beneficiaries of the research. Ryan (2002), for example, uses partial equilibrium models to calculate the economic value of the time saved in hastening welfare-enhancing trade policy changes as a result of the timely provision of policy-oriented social science research and its effective dissemination. It could also include perceptions of the peers and policy-makers about such impacts.

These products are generally sequential, with evidence becoming increasingly more difficult to gather as one moves from outputs to impacts, and with the responsibility of the research staff for documentation and evaluation decreasing and shifting to independent peer evaluators (see the continuum below):

**Figure 2. IFPRI's framework for impact evaluation**



For this framework to function, researchers are required to record output and outcome/influence indicators, which will relate directly to milestones and achievements in work plans at the beginning of the year. IFPRI staff will also record policy responses, subject to verification by independent peer impact evaluators that can be captured effectively in narratives. This is not intended as a top-down compliance approach but as a performance management approach, which could eventually become part of staff evaluation. IFPRI highlights the value of developing and maintaining suitable databases of indicators of outputs, outcomes/influences and policy responses so that they become a sustainable part of the institutional memory that is not lost as individual staff leave.

Indicators at staff level are then aggregated to the appropriate project-, programme- or institute- levels. It is important to accompany these with benchmarks in order to document the gap between goals and milestones and actual achievements. IFPRI also recognise that for a more comprehensive impact assessment, a comparison with best practices of others (i.e. with-and-without IFPRI) is desirable, much like the idea of the counterfactual.

**Were any challenges highlighted in evaluating the impact of research on policy (and any particularly relating to the impact of *environmental* research)?**

This paper highlights a number of challenges to address when trying to evaluate the impact of research on policy.

First, is the variable reliability of information from key informant interviews; some key informants will have been more closely engaged with the research than others, some may play more central roles in the development of policy, some may have more or less reliable recall of key conversations and critical events, and some will have changed posts and lost contact with the programme/policy in question before final decisions were made.

Second, is trying to find the best balance between conducting in-depth interviews with a few key informants and facilitating wider engagement through the use of written survey questionnaires: the latter is less resource intensive and enables contact with more research users, but also potentially provides less detail and credibility.

Third, is the need to look beyond the use of traditional research impact indicators such as the number of publications in leading journals and subsequent citations, as this study indicates that this measure does not correlate well with policy influence or wider impact. Indicators need to be developed that can identify less tangible impacts, such as changes in common ways of examining policy processes that lead to multiple changes in policy decisions across countries, institutions, and/or individuals. These sorts of impacts are difficult to trace and capture; this becomes increasingly difficult beyond the project level as more actors become involved, together with exponentially greater information sources and motivations for use.

Fourth, is the need to consider the duration of the impact; for how long does the research provide benefits to the policy-making process, and how long do policy changes made in response to the research remain unchanged?

Fifth, is the issue of the appropriate timing for the evaluation process; research results can set the stage for a policy change that then occurs at a much later date when the original research results are forgotten i.e. the link between the earlier research and the substantially later policy change is broken, and key informants may no longer make the link. If the evaluation is conducted too soon after the completion of the research, however, no impacts may yet be apparent.

This links to the ‘Cassandra problem’ (Ryan and Garrett, 2004); is it possible to account for and value ‘good research advice’ that is not taken? Or delays in taking the advice? It is suggested that an evaluator could calculate the opportunity costs of a ‘wrong decision’ (i.e. estimate the cost of the alternative to not taking the advice). Alternatively, so-called ‘good advice’ may have arisen from flawed research, in which case it is to the policy-makers’ benefit if they did not choose to accept it.

Finally, is the issue of attribution; how to measure the direct effects of POR and the indirect or induced effects on policy, given the diversity of additional factors that can influence the policy-making process. This may be addressed in various ways, including:

- The use of the counterfactual, though this can be challenging in itself.
- Demand-side approaches to impact evaluation (as opposed to supply-side), which uses major policy events as the starting point and works (retrospective analysis) to establish the separate influences of the many research suppliers and other factors on policy responses. Such retrospective narratives can then form the basis for quantitative estimates.
- Institutionalising the process of impact evaluation, i.e. following IFPRI’s approach in ensuring that staff and management take responsibility for recording outputs, outcomes/influences and policy responses related to their research. Independent evaluators can verify these and translate them into meaningful measures of their

impact on economic welfare, assessing to what extent any changes in welfare can be attributed to policy research institutions and their partners.

A key challenge in evaluating the impact of *environmental* research in particular is that such research is primarily likely to produce environmental benefits for society. It is often difficult to quantify such benefits as they tend to lack market values, and the development of effective environmental valuation techniques continues to prove challenging. Environmental benefits are also very hard to estimate due to the many confounding factors that have to be controlled for and the significant data-gathering requirements.

### **Information sources for case study**

CGIAR (2008) 'Impact Assessment of Policy-Oriented Research in the CGIAR: Evidence and Insights from Case Studies' *CGIAR* [online]. Available from:

[www.sciencecouncil.cgiar.org/fileadmin/user\\_upload/sciencecouncil/Impact\\_Assessment/PORIA2008\\_pr\\_4\\_F\\_1-r\\_3\\_.pdf](http://www.sciencecouncil.cgiar.org/fileadmin/user_upload/sciencecouncil/Impact_Assessment/PORIA2008_pr_4_F_1-r_3_.pdf)

Ryan, J. (2002) 'Assessing the impact of food policy research: rice trade policies in Vietnam' *Food Policy* 27: 1-29

Ryan, JG. And Garrett, JL. (2004) 'The impact of economic policy research: Lessons on attribution and evaluation from IFPRI' in Stone, D. and Maxwell, S. (eds) (2004) '*Global Knowledge Networks and International Development: Bridges across boundaries*'. Routledge, London.

Zilberman, D. and Waibel, H. (2007) 'Advances in impact assessment of natural resources management research' *Quarterly Journal of International Agriculture* 46(4): 395-420

Email correspondence with James Garrett and Jim Ryan.

## **Case Study 4: Defra/Environment Agency FCERM Programme**

**Name and background of organisation:** Joint Defra/Environment Agency for England and Wales.

Defra is the government department in the UK with responsibility for securing a healthy natural environment by dealing with environmental risks and promoting a sustainable, low-carbon and resource-efficient economy. Defra also has responsibility ensuring a thriving farming sector and a sustainable, healthy and secure food supply. Defra funds the regulatory function of the environmental regulator, the Environment Agency.

The Environment Agency is the environmental regulator for England and Wales and is responsible for a broad range of environmental issues including pollution to air, land and water, waste disposal, water resources, flooding, climate change and environmental monitoring.

### **What is the background to the study?**

The Flood and Coastal Erosion Risk Management R&D Programme (FCERM) is a thematic research and development programme, guided and peer-reviewed by experts in relevant fields. It focuses on both the development of relevant evidence and the appropriate delivery of research outputs to end-users.

The Joint Programme aims to help inform policy development and ensure that flood and coastal erosion risk management measures are delivered in a sustainable and cost effective manner. It provides key evidence, information, and tools and techniques required to develop policy and strategy for flood and coastal erosion risk management. It also needs to have the ability to:

- Understand and assess flood and coastal erosion risks and the processes by which these risks arise;
- Manage the assets used in flood and coastal erosion risk management sustainably; and
- Prepare for and manage flood incidents effectively.

The Programme aims to address the needs of all flood and coastal defence authorities and bridge the gap between work carried out by others (basic scientific research) and the development of operational systems.

The 2005 Joint Programme Review (Defra 2005) found that the Programme has generally been a success and has delivered a wide range of useful outputs, aimed at both practitioners and researchers, but also identified a number of areas requiring change including in the evaluation of research:

*‘We believe that there is an urgent need for better evaluation of projects and themes and recommend that this is done, on an annual cycle, using a common reporting and assessment system’.*

**Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

This approach is primarily derived from practical need rather than theory, and is a system-based approach.

**What evaluation approach was used to assess policy impact?**

The approach adopted for this evaluation acknowledges that identifying the costs and benefits arising from research can be difficult, especially when they arise several years after the completion of a study. The FCERM programme has a suite of measures to provide key information on benefits, with proportionate effort required to measure and monitor. The first stage focuses on the monitoring of uptake of outputs by a targeted end user community. A further stage is aimed at assessing what wider use is being made of the research in the longer term.

The evaluation process is extensively embedded into the management structures of the programme. Each individual research project has a Defra or Environment Agency project officer who is involved in the project from conception to post-project evaluation. The responsibility for delivering project outputs lies with the Project Executive. The Project Officer/Manager works with the Project Sponsor and the Project Board/Steering Group (if in place) in order to provide evidence that the project has successfully achieved and delivered in line with business needs.

***Project benefits.*** A central component of the evaluation relies on the SMART Benefits Delivery Plan which is developed for each project at project inception by the project officer. It is then the Project Executive who is responsible for ensuring that the project meets this benefits plan. The identified user of the research is responsible for steering the outputs to the business during the course of the project and beyond. Theme Managers remain in closely contact with business users in order to monitor benefits realisation. The following questions are considered in monitoring and evaluating benefit realisation.

1. Have all business changes required to exploit the products been introduced into the organisation?
2. Are the benefits as set out in the Benefits Plan being realised and if not are plans in place to remedy this?
3. Is benefit delivery actively reported on as a component of performance management?
4. Are lessons learned being captured and reported?
5. Are arrangements in place for future Benefit Realisation Reviews agreed?

*Programme benefits* are assessed by comparing Programme/Project deliverables against the Programme Rationale & Objectives Statement. A Programme Benefits Realisation Plan is maintained, including a complete description of benefit, time scales and evidence on benefits realisation. Whilst Project Officers/Managers will ensure the appropriate links are managed and maintained during the project, it is the Theme Manager(s) who continue to monitor uptake and support the Programme Manager(s) in maintaining the Programme Benefits Realisation Plan. Alongside the Benefits Plan, the Programme Manager(s) maintains a log of problems and successes. The Benefits Realisation Plan is reviewed by the JPB (Joint Programme Management Team) annually. The following questions are considered in monitoring and evaluating benefit realisation.

1. Developing the FCERM knowledge base:
  - With reference to the Joint Programme Rationale & Objectives statements, what was/were the key ‘answer(s)’ provided by the research?
  - Has the project identified important ‘knowledge gaps’ and/or the need for, further research?
2. Policy, Process and operational outcomes:
  - What are the key outcomes of the research outputs, for example, in Policy development, Process management or Operational good practice (outline in one paragraph)?
3. Cost savings and resources:
  - Has the research led to any cost savings that can be (reasonably) readily quantified, for example in staff time or delivery/monitoring costs?
  - Has the research been able to lever further resources into the total amount for flood related R&D (e.g. through contributing to R&D programmes of other funders)?

**Were any challenges highlighted in assessing the impact of the organisations work (and particularly those of *environmental* research)?**

The 2005 review recommended that the Programme Board and the Programme Manager adopt an integrated model for: (Problem identification; Research and development; Targeted dissemination; Implementation, and Evaluated use) and highlighted an urgent need for better evaluation of projects and themes and recommended that this be done, on an annual cycle, using a common reporting and assessment system. Since then, however, systems and benchmarks have not been developed to allow Programme and theme evaluation to happen as a matter of routine. There has been insufficient monitoring and evaluation of the Programme, so that it is not possible to identify clearly the benefits that are obtained from FCERM and relate these to the investment in specific R&D projects.

It was unclear how successful the programme was at helping to deliver an increase in the effectiveness and efficiency of flood and coastal erosion risk management.

**Information source(s) for the case study**

Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme:  
Programme Definition, December 2006. Details from:  
[www.defra.gov.uk/environ/fcd/research/RandDProgCon/PDDDecember06.pdf](http://www.defra.gov.uk/environ/fcd/research/RandDProgCon/PDDDecember06.pdf)

Defra 2005, An Independent Review of the Defra/EA Research and Development Joint Programme in Flood and Coastal Erosion Risk Management by Edmund Penning-Rowsell; Peter Bye; Charlie Rickard; Ian Townend; Andrew Watkinson, Produced: June 2005.

## **Case Study 5: The Finnish Environment Institute (SYKE)**

**Name and background of organisation:** The Finnish Environment Institute (SYKE).

SYKE is Finland's national centre for environmental research and development, and is part of the environmental administration of the government. SYKE was established in 1995 to work at the interface of science and decision-making, focusing research on changes in the environment and how to control these changes. SYKE's overall research programme is divided into a number of individual programmes on: global change, production and consumption, contaminants and risks, integrated river basin management, protection of the Baltic Sea, biodiversity, and environmental policy.

The main 'customers' of SYKE's work are the various ministries, especially the Ministry of the Environment and the Ministry of Agriculture and Forestry as they supervise SKYE and provide 60 per cent of their funding. The Ministry of the Environment does not have its own scientists and uses SYKE as its scientific arm, developing the scientific knowledge needed to support policy-making, providing scientific advice to policy-making processes, and supporting the Ministry in international negotiations.

Each year, SYKE negotiates with these two ministries for funding for the next year, which results in a formal agreement on what they will do for them. The remaining 40 per cent of SYKE's budget is obtained from national and international research and development donors such as the Academy of Finland and the DG Research

### **What is the background to this study?**

Evaluation specifically of the uptake of SYKE's work is included within its wider evaluation strategy. The aim of this aspect of evaluation is to enable SYKE to identify how they can best increase the impact of their work on society.

### **Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

SYKE bases its theoretical approach to evaluation on a triangulation of methodologies and combines various criteria. Intervention theory has appeared useful. In this, the person in charge of an intervention (whether a research manager, research donor or a decision maker) illustrates the theory he had when the intervention was planned and this is then analysed or, if it is a question of ex post evaluation, assessed against the original intervention theory. Intervention theories are held by many people for the same intervention and it is fruitful to compare them. This method is mainly used for policy evaluation but functions also for evaluation of research uptake as an intervention.

Aside from these two theories, the approaches used for evaluating uptake of research have primarily emerged from practical requirements and constraints. All of SYKE's research aims to be policy-relevant, rather than blue skies research. In some ways this facilitates the evaluation of its uptake as all research tends to be planned in collaboration with its

intended policy users, in order to be used in specific policy processes, i.e. the research is problem-oriented. The evaluation processes therefore make efforts to identify how the research has been framed and with whom, and the number of research projects which are interdisciplinary (as problem-oriented research tends to require interdisciplinary approaches).

### **What evaluation approach is used to assess policy impact?**

SYKE conducts several layers of general evaluation, each of which incorporates some effort to evaluate the policy *uptake* of their research, largely through conducting interviews/surveys with research ‘customers’ and the researchers themselves.

#### *a) Top-level, ten-year independent evaluation:*

SYKE is evaluated every ten years by an independent international evaluation panel, the latest of which was undertaken from September to December 2008. As part of this evaluation, 200 potential users of SYKE’s work were interviewed, either face-to-face or via email, to establish for example how they use the research produced. The interview consisted primarily of open questions, in which interviewees were asked, among other things, to highlight where they have used the research outputs, e.g. which policy processes the research has contributed to, which documents the research has been cited in, whether the research has been used to justify particular stances in policy discussions etc. They are *not* asked to rate or value the contribution of research, or its impact.

#### *b) Four year strategic evaluation*

This evaluation is conducted internally, when reviews are undertaken of SYKE’s strategic goals. The self evaluation conducted and cross-checked by managers of different research programmes in different divisions of SYKE is followed by stakeholder interviews/surveys with external customers, with those in the ministries. The findings from these strategic evaluations can be considered alongside the findings of the ten-year evaluation, to consider whether and how to change current research directions and approaches to better align with the needs of key stakeholders.

#### *c) Annual evaluation*

Internal evaluations are also conducted each year through annual reporting, including both an evaluation of the previous year’s results, and of their impact and effectiveness in supporting customers. This draws on the ‘balanced scorecard’. For some projects, the SYKE project leader and the supervisor in the Ministry discuss the success of the project including elements, which could be referred to as criteria of the uptake of them in policy making:

- Organisation of the project
- Budget issues
- Timetable issues
- How easy it was to agree things
- How well the project met its objectives

- Research methods
- Networking aspects
- Effect on stakeholders
- Dissemination: the Internet, seminars, other researchers, how well publications succeeded.
- Co-operation with other projects and stakeholders.

The project leader and the Ministry supervisor score independently, in order to make the evaluations more trustworthy. There is generally a good match between the scores, and where there are particularly high or low or divergent scores the project leader is interviewed (recognising that more can often be learnt from the most extreme cases).

On selected thematic areas, SYKE representatives also attend an annual meeting with the ministries to explore the emerging knowledge needs and research, to consider future priorities for SYKE research such that it aligns with government needs, and also to discuss how the ministry has been using the research produced.

*d) Individual self-evaluation*

Yearly negotiations are conducted with each SYKE staff member to assess how they have developed their research, progress made towards achieving research goals, the extent to which their research has been used (e.g. looking at citations in particular policy documents/reports), by whom and how it has influenced decision-making. These self-evaluations can feed into the annual evaluation process.

**Were any challenges highlighted in assessing the impact of the organisation's work (and particularly of *environmental* research)?**

A number of generic challenges were identified when trying to evaluate the uptake of research:

Firstly, whilst not applicable to SYKE's research, the difficulty in assessing the uptake of more basic or blue skies research was mentioned, as research planned and executed without the input of or initial demand from policy-makers is rarely taken up directly or immediately by policy-makers.

Secondly, the issue of attribution was also raised; much of SYKE's research is collaborative and it can be difficult to establish SYKE's individual contribution to a policy change. Furthermore, research can be used in many different ways and at several different levels, which complicates efforts to trace uptake. However, it can be possible to trace uptake in some cases, e.g. the debate in Prime Minister's Office and the Parliament on energy and climate strategy used SYKE's work and the OECD's country study on Finland's environmental governance had direct references to SYKE's work. Similarly, the policy evaluation conducted by SYKE of the Finnish National Biodiversity Strategy fed

into the development of the new National Biodiversity Strategy, with particular use made in certain areas of Finland which are conflict-prone from a biodiversity perspective.

Thirdly, was the time-lag issue; the length of time elapsed between the completion of research and its uptake can be highly variable. Evaluations may therefore need to be conducted at several different time intervals.

*Environmental* research findings are often accompanied by considerable uncertainties (e.g. climate change research), which can deter policy uptake/impact. Impact evaluations may therefore need to consider how these uncertainties have been explained to the policy community, to fully understand how uncertainties may have compromised the potential uptake of the research. In addition, policy-makers need to be honest in how they communicate uncertainties to ensure success of the policy in question.

#### **Information source(s) for case study**

Interview with Eeva Furman (Programme Manager, Environmental Policy Research Programme) and discussion with Dr. John Holmes.

Reference to the appendices of: Holmes, J. and Savgard, J. (2008) 'Dissemination and implementation of environmental research' SKEP Work Package 4 [online]. Available from: [http://www.skep-era.net/site/files/WP4\\_final%20report.pdf](http://www.skep-era.net/site/files/WP4_final%20report.pdf)

## Case Study 6: Foundation for Research, Science and Technology, New Zealand

### **Name and background of organisation**

Foundation for Research, Science and Technology (FRST), New Zealand.

FRST was established by an act of government in 1990 to invest in research, science and technology (RS&T) for the benefit of New Zealand. Selected research could improve the standard of living of New Zealanders by improving environmental or social conditions, or it could boost the economy by delivering a new niche product or service.

FRST is a Crown entity governed by a board appointed by the Minister of Research, Science and Technology. The Foundation is the largest public sector investment agent for the Government's investment in research, science and technology (called Vote RS&T), investing three quarters of the money allocated by the Government to RS&T.

### **What is the background to this study?**

In 2002/03, FRST conducted a survey to evaluate the outcomes of their natural ecosystems research, which aims to enhance the management and protection of New Zealand's natural ecosystems. This survey was intended to feed into a wider review of its natural ecosystems research, with a view to reinvesting in that area in 2004/05. The evaluation was focused at an overview level, and there was no intention to use the findings to influence investment decisions for individual research programmes.

### **Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

The evaluation methods chosen largely arose out of practical need, although there was a loose conceptual framework informing the development of their approach. They wished to collect evidence from a variety of people reflecting the differing perspectives of researchers and research users, and they aimed to use a combination of tools to collect quantitative and qualitative evidence (such that the quantitative evidence could provide data to calibrate the qualitative evidence).

They felt that one aspect of their approach was particularly innovative; using independent consultants to select the evaluation participants and in particular, focus on selecting 'key players' or 'information nodes' rather than drawing from the general pool of users or a pool identified by researchers. The consultants used a matrix template with the type of research on one axis (freshwater, marine, land etc) and type of user organisation (national agency, local authority, conservation group etc) on the other axis to ensure they had good coverage of users.

### **What evaluation approach was used to assess outcomes?**

The evaluation consisted of two parts: a user survey and case studies of ecosystem research programmes, with the main emphasis on the user survey.

### The Case Studies

The evaluation unit undertook 12 case studies in the ecosystem research area as part of a broader evaluation between 2001 and 2003. These provide qualitative information about the nature and impact of benefits arising from the FRST-funded research. The primary perspective in the case studies is that of the researchers, but also included corroboration with users identified as beneficiaries of the research results.

### The User Survey

The aim of the survey was to obtain information on the natural ecosystem research funded by the Foundation, from the perspectives of people involved in the management and/or protection of New Zealand ecosystems. The survey sought views on: the benefits of natural ecosystem research; relationships with researchers; factors affecting the implementation of research results.

The user survey had two strands:

Firstly, a short web-based questionnaire - intended to reach a large number and diversity of users (not just those identified by the researchers themselves, but instead users who could potentially benefit from the research, as identified by independent consultants). The potential user population ranged from central and local government policy and operational agencies, to community groups. Consultants were also identified as important users to be represented in the survey due to their role in supplying management and policy advice to regional and local councils.

Secondly, an in-depth interview-based survey - this was carried out with fewer respondents (32 face-to-face interviews in total), in order to explore the outcomes in more depth. These users were selected on the basis that they should be more familiar with the RS&T system, were well connected within that system, and to some degree played an 'information node' role. A mixture of open and closed questions were asked, broadly relating to:

- Background information about the user, e.g. organisation, role, activities related to natural ecosystems, main sources of knowledge used etc.
- Awareness of FRST-funded research – users were asked to list and briefly describe projects they had heard of. A list of projects was available as a prompt if required.
- Number and purpose of relationships with FRST researchers and factors affecting the relationships.
- Nature of benefits obtained from FRST research.
- Rating of usefulness of FRST research for achieving benefit (1 = not useful, 7 = extremely useful).
- Rating of impact of each benefit/outcome on their organisation/work (low, moderate, high).
- Barriers or gaps associated with FRST research.
- Other general comments on relevance/value of FRST research.

Information from the web-survey was intended, to some extent, to corroborate the more detailed but fewer in number interview reports.

In total, 176 people were asked to participate in either a web or interview survey, with an overall response rate of 72 per cent. 47 per cent overall respondents were in a central government agency, and 32 per cent in a local or regional government agency.

Following this FRST evaluation, a wider evaluation of environmental research was undertaken by the Ministry of Research Science and Technology using a similar methodology to this one (as outlined in:

[www.morst.govt.nz/Documents/publications/evaluations/Environmental-Output-Class-Evaluation-Report.pdf](http://www.morst.govt.nz/Documents/publications/evaluations/Environmental-Output-Class-Evaluation-Report.pdf)).

This included some case studies, face-to-face interviews with regional councils and also an international comparison of institutional and funding arrangements in six different OECD countries. One of the outcomes of that evaluation was a new government scheme called 'EnviroLink' which was set up to facilitate uptake of environmental research by regional councils. This scheme was evaluated after a 16 month trial period (more info available from:

[www.morst.govt.nz/Documents/publications/evaluations/EnviroLink%20Evaluation.pdf](http://www.morst.govt.nz/Documents/publications/evaluations/EnviroLink%20Evaluation.pdf)).

It was suggested that if repeated, a more structured approach to synthesising the case studies and structured interviews would be adopted, and efforts would be made to try to adopt a research value mapping methodology. This would involve identifying the most common routes by which environmental research is transferred to and adopted by users, and the key success factors involved at various stages. An example of the use of this methodology in a non-environmental context may be found at:

[www.frst.govt.nz/files/20041201 ICT Value Mapping Report Dec2004.pdf](http://www.frst.govt.nz/files/20041201 ICT Value Mapping Report Dec2004.pdf)).

**Were any challenges highlighted in assessing the impact of the organisation's work (and particularly of *environmental* research)?**

Two general challenges were highlighted:

1. *The Attribution Issue:* The Foundation research programmes are not able to claim sole credit for outcomes that were achieved in association with other government agencies or groups, but neither are they fully responsible for factors that affect the successful implementation of research results, which may be caused by system-wide deficiencies.
2. *Diversity of the User Population:* With such a large and diverse potential user population, not all the research will be of value to everyone, which can be difficult to account for when interpreting the findings of the evaluation.

It was felt that the potential user population is particularly diverse in the case of *environmental* research in New Zealand. There are a number of different agencies responsible for environmental policy and management of public and private land. These range from national agencies such as the Ministry for Environment and Department of Conservation to regional and local authorities. FRST addressed this challenge by hiring independent environmental consultants to pinpoint some of the key people with strong networks and/or influence to participate in this evaluation.

On a similar note, it was also noted that the pathways to achieving outcomes from environmental research were more complicated than other types of research. Clearly, the dissemination and implementation of research results cannot be regarded as a linear process, but there was a feeling that the process was less linear for environmental research compared with research targeted at existing manufacturing industry, for example. FRST felt that they did not really address this fully except by including face-to-face interviews and collecting detailed qualitative data on outcomes and important success factors, such as collaborations.

Another challenge for environmental research is that intangible outcomes are considered to be very valuable and these are more difficult to measure. For example, users regarded new knowledge which might lead to a change in awareness or understanding of an environmental problem as being as important to the development of a new management tool. Similarly, some users thought that maintaining national capability in important areas such as insurance policy was one of the key outcomes of basic environmental research. This is quite different to many other research areas looked at where the main outcomes are more often tangible benefits (e.g. manufacturing research). FRST addressed this challenge by ensuring they included questions to capture these benefits and their impact in their web surveys and face-to-face interviews. They also followed the evaluation with a cost benefit study focused at a type of environmental research programme. This was partially intended to address sceptics' views that the only outcomes from environmental research were 'soft' rather than hard economic results.

Finally, a further challenge was to narrow the scope of the evaluation to areas that were sufficiently alike as to enable the evaluators to draw appropriate conclusions about the benefits and challenges faced by researchers. For example, the implementation pathways might be quite different for research targeted at the land environment as compared to marine. Some arbitrary decisions had to be made to exclude some research areas.

#### **Information sources for case study**

Report: FRST (2003) 'Natural Ecosystem Research Evaluation' [online]. Available from: [www.frst.govt.nz/files/Natural Ecosystem Research Evaluation.pdf](http://www.frst.govt.nz/files/Natural_Ecosystem_Research_Evaluation.pdf)

Email discussion with key report author, Michele Morris.

## **Case Study 7: Environment Agency for England and Wales, Post-Project Appraisals**

**Name of organisation:** Environment Agency for England and Wales

This case study concerns a set of five post-project appraisals carried out by the Environment Agency for England and Wales in 2006. The Environment Agency is the environmental regulator for England and Wales and is responsible for a broad range of environmental issues including pollution to air, land and water, waste disposal, water resources, flooding, climate change and environmental monitoring. Its research programme is intended to generate the information, data, and models required to support policy-making and operational decision taking. The budget for research projects is around €10 million annually.

Most research projects are commissioned externally with consultancies, research institutes and universities; just a few are carried out in-house. Typically, around 300 research projects will be 'live' at any one time. The Agency's Science Department is responsible for the planning and management of the research: a project support officer from the Science Department is appointed to each research project. 'Business customers' for the research sit within the Head Office Policy functions, the 26 Areas where many of the day-to-day operational decisions are taken, and the 'Process' teams which sit between Head Office Policy and the Areas translating policy into procedures and ensuring consistency of approach across the Agency.

A 'Cradle to Grave' (C2G) project management system endeavours to ensure that research projects are well-focused on delivering 'business' benefits from their inception through to the uptake and application of their outputs. Within this system, each research project will have a project sponsor (a senior manager from Policy, Process or Operations who oversees the project and is responsible for ensuring that any business obstacles to uptake and application are overcome), a project executive (responsible to the project sponsor for ensuring that the intended benefits from the project are realised), and a business user (the representative of the business unit that will use the outputs of the project).

### **What is the background to this study?**

Historically, around 5 to 10 post-project appraisals had been carried out each year, which aimed to evaluate an indicative sample of the research programme. The focus of these evaluations was the effectiveness of project management processes and their adherence to Agency procedures. They were carried out by an external contractor, and based on review of project documentation and interviews with the involved staff.

The context for these post-project appraisals was an increasing emphasis on realising the intended benefits of research projects. Also, there was a concern that the impacts and value of the Agency's research programme was not as visible to its 'customers' in policy, process and operations as it could be. Both delivering, and being seen to deliver, benefits

were important in making the case for maintaining the Agency's investment in its research programme in the face of increasing competition for reducing budgets. As well as continuing to examine project management process issues, the post-project appraisals therefore also endeavoured to evaluate uptake and impact.

A selection of five research projects which had been completed in 2004 and 2005 were chosen to be evaluated:

1. Development of flood warning management system
2. Oestrogenic effects in wild fish post spawning exposure to oestrogenic effluents
3. Strategic research project into the natural attenuation of groundwater pollutants
4. Reworking intermediate level radioactive waste packages
5. Developing soil screening values for use in ecological risk assessment.

They were selected to represent a range of tendering processes (competitively tendered, collaborative, and single-tender), a large 'umbrella' project, and a project carried out in-house.

**Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

The approach appears to have evolved through practical need rather than from a firm theoretical foundation.

**What evaluation approach was used to assess policy impact?**

It was decided to carry out the evaluations using a workshop format, rather than to continue with the previous practice based on review of documentation and interviews. It was felt that this should provide a better opportunity to draw out and explore issues. A two hour workshop was held for each of the five projects. The identified purpose was to identify lessons that could be learned to improve future project management practices, and in particular to ensure that the Agency derives maximum value from its research programme. It was considered important to foster an atmosphere in the workshops of 'learning together', and was expressly not about 'holding to account'.

The workshops involved between three and seven people and were facilitated by the staff members from Science Department responsible for the evaluation initiative. Attendees were staff involved in ensuring delivery of the projects' products and implementing them to deliver business benefits. On a practical note, it was found to be important to get invitations to the workshops out early so that key players were represented.

The workshop itself was divided into three broad elements. Following a brief review of the aims of the workshop, a brief brainstorming session was held to identify the strengths and weaknesses of the project. This was followed by more detailed reviews of 'benefits delivery' and 'project processes'. A concluding section looked at lessons learned and recommendations for future projects. A report of the workshop was produced using a

standard format (see Figure 3 below). Attendees were given the opportunity to comment on this report in draft to ensure it was a fair reflection of opinions expressed in the workshop.

A report on the findings of the five workshops was considered by the Agency's Project Appraisal Board (which oversees the conduct of the research programme and its quality assurance) in May 2006. It, in turn, made recommendations to the senior bodies responsible for the programme: the Science Management Executive Team and the Science Programme Board.

A revised approach to research project evaluation is currently being developed.

**Were any challenges highlighted in assessing the impact of the organisation's work (and particularly of *environmental* research)?**

Some reluctance to participate in the evaluation process was encountered in some cases. This was particularly so for more senior staff on the customer side, notably the project sponsors and project executives. It was considered that post-project appraisal should be seen as an integral part of the roles of staff involved with research projects, whether on the project manager or customer side. Where staff did attend the workshops they found them to be enjoyable and beneficial: real lessons were learned that may have been missed otherwise. There was a good level of openness and frankness.

With one exception, the umbrella project on natural attenuation of groundwater pollution, the evaluations of impact were unable to identify a monetary value to the projects' contributions to the intended environmental outcomes and/or costs avoided. It proved to be easier to evaluate uptake than impact.

The timing of the evaluation was problematic. Conducting it soon after research project closure means that it is more likely that the relevant staff will still be available and will have a sufficiently detailed memory of events. However, sufficient time may not have elapsed for uptake and impact to be apparent. Conversely, holding the evaluation later runs the risk that staff will have moved on and/or may not have sufficient recall of events. It was suggested that a two-stage evaluation process may be appropriate: the first stage being held soon after research project closure and relying on documentation and a limited number of interviews; a second stage adopting the workshop format and being held when the benefits of the project may be expected to be realised.

A distinctive characteristic of much environmental research is that it is interdisciplinary, and increasingly brings together the natural and social sciences and evaluation approach needs therefore to reflect this.

**Information source(s) for case study**

Personal experience of project team member.



**Figure 3. Post-project Appraisal – Record of workshop (Template)**

Project No:  
 Project Title:  
 Date:  
 Attendees:

*Section 1: Introduction*

Introductions  
 Background to the need for PPA's and usage of reports etc

*Section 2: Overview*

<b>Question</b>	
What were the projects key strengths?	
What were the projects key weaknesses?	
What were the causes of the strengths?	
What were the causes of the weaknesses?	
What are the main lessons identified at this stage?	

*Section 3: Benefits realisation*

<b>Question</b>	
Clarify the need for the project, and how it links to science strategy, business needs etc	
Are copies of any output reports etc available?	
Was a benefits plan produced?	
Were the objectives / planned benefits SMART?	
Have / are delivery of benefit's being monitored?	
Have all business changes required to exploit the products been introduced into the organisation?	
Are business users using the products as planned?	
Are communications on any change operating efficiently?	
Are any training materials / programmes up to date?	
Are any revised policies, procedures and working practices being adhered to?	
Have any obsolete practices / assets been managed out of the business?	
Overall how do delivered benefits compare to those that were planned and is there demonstrable evidence of delivery?	
Reasons for any discrepancy in scale or timing of benefits?	
Is there any plan in place for further measurement / review of benefits and continual improvement (Links to KPIS / business planning/programme management)?	
Key strengths and weakness	
Identification of lessons	

*Section 4: Project process*

<b>Question</b>	<b>Document review</b>	<b>Workshop</b>
Was AMS Project Management Procedure referred to?		
Is there existing / accessible project paperwork?		
Was an Outline Risk Assessment produced?		
Was relevant Business Justification documentation produced? (Form A, Business Case)		
Were all other options fully considered and costs / benefits of each assessed?		
How was this project prioritised / assessed against others within the programme?		
Is there evidence of appropriate FsoD approval?		
Is there a complete project plan comprising: Schedule of tasks, time resource; Risk Register; Benefit Plan; Budget Plan / expenditure records; Procurement Plan; Product Description; Governance Structure; Change Log.		
Were regular Review Stages undertaken? Did they fulfil their objectives and how?		
Were stakeholders (internal/external) identified at the start of project? What communications were planned and undertaken?		
Were risks successfully managed, did any arise that had not been planned for and how were these dealt with?		
How did the project team work together?		
<i>If contractor:</i> Did the KPI process work, what happened with measures? Was the relationship positive / negative? Did we / they manage roles effectively? Did the contractor contribute to an effective project? Did the contractor bring any innovation? Did they meet their commitments under the contract? Comments contractor may have on working with EA and our process?		
Is a peer review of the work undertaken appropriate to this project? If so has it been undertaken / planned?		
In general was the project well planned and controlled throughout?		
What were the critical success factors of the project and comparison of delivery against these?		
Key strengths and weaknesses		
Identification of lessons		

*Section 5: Conclusions*

<b>Question</b>	
Review of lessons identified	
Final observations	
What happens next and any comments on the PPA process	

## Case Study 8: Swedish Environmental Protection Agency

**Name and background of organisation:** Swedish Environmental Protection Agency (SwEPA), Sweden.

The main task of the Swedish Environmental Protection Agency (EPA) is to ensure that environmental policy decisions are implemented. The Agency supports the Ministry of Sustainable Development on national and international environmental issues by providing background material and scientific advice. The reliance of the Ministry on EPA staff is reflected in their respective numbers of employees: the Ministry employs around 100 people, whereas the EPA 550.

The Swedish Parliament, the Riksdag, adopted *16 environmental goals*<sup>4</sup> aimed at tackling the country's major environmental issues by 2020. These environmental quality objectives provide the framework for all environmental work in Sweden. The Swedish Environmental Protection Agency has overall responsibility for achieving these environmental objectives. Their efforts are guided by a fixed structure of regular monitoring, evaluations and proposals for changes.

Five years ago the EPA introduced a system of *integrated responsibility* for research. Consequently, the four operational departments of EPA have the responsibility for defining research needs and monitoring the individual research programmes. One of the departments, the Research Secretariat, is responsible for the overall planning and coordination of the programme.

### What is the background to this study?

The evaluation of environmental research is an integral part of the planning cycle of the EPA. The planning cycle comprises:

1. Identification of research needs;
2. Mid-term evaluation (at the programme and subtopic level);
3. Process-evaluation (budget, personnel, communication inside and outside the program, products, timescale etc);
4. An evaluation of achievement of the national environmental goals e.g. identification of results which have contributed to the goals above.( made half a year after end of project/program);
5. Scientific evaluation - two years after end-of-project a bibliometric analysis is made (identification of publication in relation to world mean, networking status of the researches in relation to subject).

In the first stage of the planning cycle the end users (national policy makers as well as policy makers on the local level) are involved in defining the research needs. Their

---

<sup>4</sup> Sweden's environmental objectives in brief and a summary of the Swedish environmental objectives council's evaluation, 2008 (<http://www.naturvardsverket.se/Documents/publikationer/978-91-620-8324-3.pdf>)

involvement in the early stages of the planning cycle is believed to enhance the research relevance and uptake of the results. The EPA has the obligation to report to the government each year on the progress that is made regarding environmental objectives, and it is also important that the EPA is able to learn from the planning and management of previous research projects and to take the lessons learnt into account in the next planning cycle.

The three main objectives of the research evaluation are to assess the quality, relevance and the management process of the research.

**1. Quality.** The main instrument for evaluating the quality of research is a bibliometric analysis (e.g. Sandström and Sandström, 2009<sup>5</sup>). A high number of citations is a measurement of the scientific quality of the research and consequently gives policy makers confidence in its use. The bibliometric analysis is based on available methods from literature. By means of international comparison, corrections are made for the research area (e.g. the number of citations are typically higher in chemistry compared to the social sciences).

**2. Relevance.** The relevance of environmental research is checked against the 16 national environmental objectives by assessing the impact of the research on the environment. There is a handbook available in Swedish on monitoring. When measuring this impact the speed of uptake is an important factor to be considered. For example, in Sweden research was already being performed ten years ago on lifestyle changes for energy saving and climate change mitigation. At that time there was not much demand for these research results whereas currently this is considered to be a 'hot topic'. In contrast, game management research immediately received a lot of attention. The effects of game management research on the environment are furthermore easier to measure since it is a relatively small research field and only a limited number of actors are involved. The impact on water pollution (nitrate and phosphorus) is more difficult to measure. Here a method for trend analysis is developed and is now incorporated as a routine method.

**3. Management process.** Evaluation of the research management process involves amongst others an assessment of the budget, research goals, deployment of researchers, timing and the interaction between the research group and stakeholders.

### **Have any particular theories or conceptual frameworks informed the development of the impact evaluation approach used?**

The approach adopted is largely based on practical considerations and requirements.

### **What evaluation approaches were used to assess policy impact?**

Two different approaches have generally been used:

---

<sup>5</sup> Erik Sandström and Ulf Sandström, Meeting the micro-level challenges: bibliometrics at the individual level, Paper to the 12th Conference on Scientometrics and Informetrics July 14-17, 2009 , Rio de Janeiro, Brazil

**1. Ad-Hoc approach.** The uptake of environmental research into policy is not systematically measured and evaluated. As mentioned above, the standard evaluation approach includes the assessment of the quality, relevance and management process. In addition, the EPA conducts different ad-hoc evaluations for different research projects. For example, the employment of a doctorandus (measurement of human capital competence) can be checked, a review by international experts (e.g. Baltic Sea eutrophication research) can be executed after the project is finished or stakeholders can be interviewed.

Sometimes, the uptake of the research results by policy is easily identified, for example when the results are directly used in EU-negotiations (e.g. research on eutrophication for the Water Framework Directive). Also, when the research field is rather small with a limited number of actors (e.g. game management research in Sweden) it is easier to evaluate whether the research has had an impact even on the local policy level by means of interviews.

**2. Backwards check from environmental goals.** One of the instruments the EPA uses in its planning cycle is the construction of a diagram in order to assess why a particular environmental goal is not achieved or not likely to be achieved. An example of one of these diagrams (related to the environmental goal of a non-toxic environment) is illustrated in Figure 4 below.

The construction of the diagram starts at the environmental goal and builds backwards towards concrete reasons for not reaching the environmental goal and the actors that are involved. The diagram does not explicitly allow for evaluation of uptake of environmental research but it does provide insight on this matter. For example, one of the reasons for not achieving a non-toxic environment is a lack of knowledge in public bodies on cocktail effects and bioaccumulation. Given the nature of the decision making process, there may be other (political) reasons for not achieving environmental goals, and various barriers to the uptake of existing knowledge by policy. The projects listed on the left-hand side of the diagram are defined on the basis of the research needs identified while constructing the diagram.

The process of constructing the diagram is led by an external consultant who provides an objective view on the matter and ensures that the process is more likely to be carried through than an exclusively internal process. Policy makers (end-users of research) are involved in the construction of the diagram.

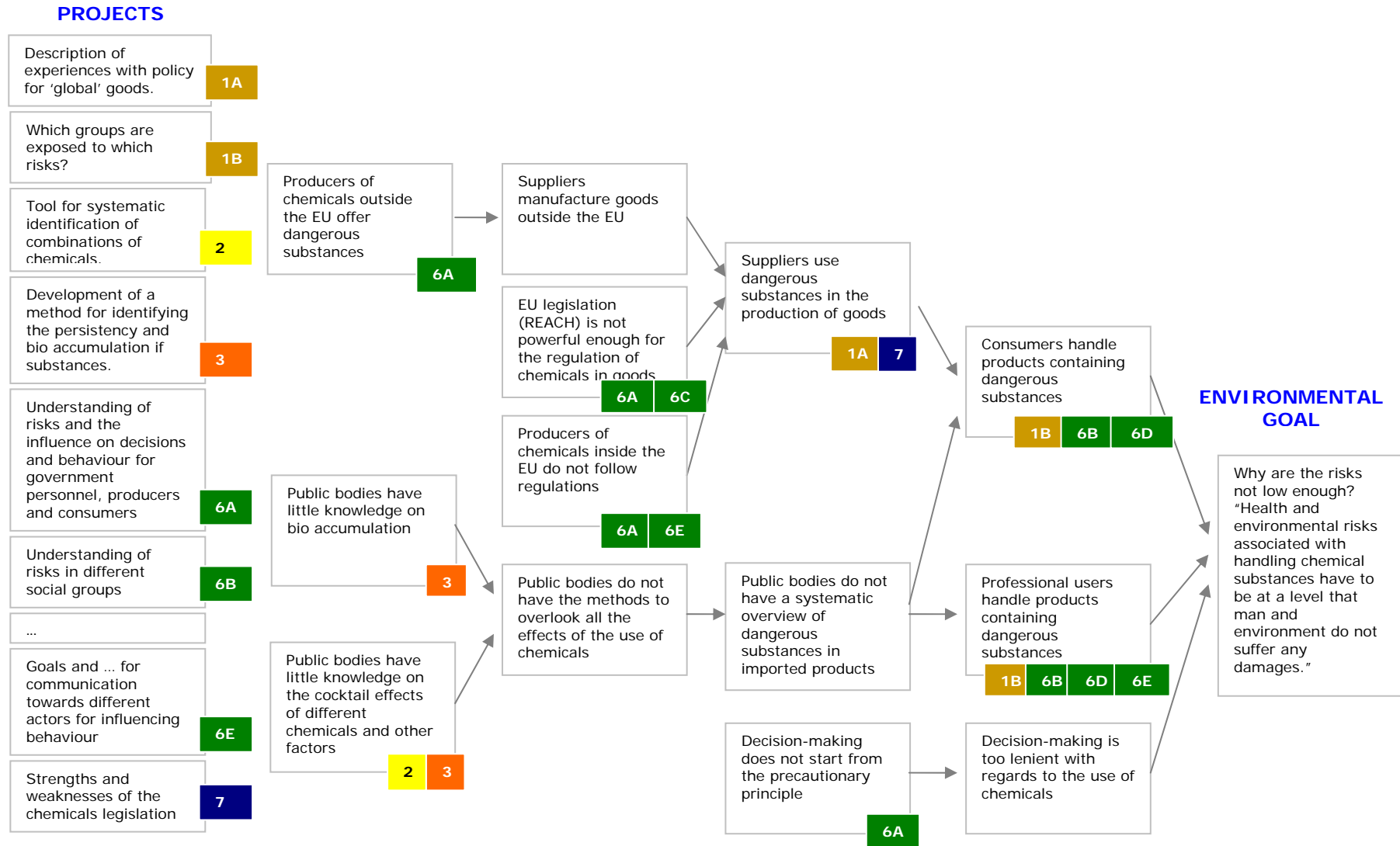
**Were any challenges highlighted in assessing the impact of the organisations work (and particularly of environmental research)?**

In the past it was difficult to assess the impact on the environment. Since 2003 the EPA uses revised monitoring and sampling methods aimed at generating enough data in order to perform trend analysis.

**Information source(s) for case study**

This case study is mainly based on an interview with Erik Fellenius (Director of the Research Secretariat) and Catarina Johansson (responsible for research evaluation).

**Figure 4: Illustration of diagram identifying possible links between research and the need for change**



## Case Study 9: Land and Water Australia

**Name and background of organisation:** Australian Government, Land and Water Australia (LWA).

LWA is an Australian Government statutory Research and Development Corporation, whose research portfolio covers sustainable agriculture and natural resource management (NRM). Its focus is on public investment in applied R&D to ‘develop and implement more sustainable systems of land and water management’ though it recognises that many of the lessons it has learned from doing this are applicable in other domains.

This is not a case study about a particular evaluation of research/policy links, but about an approach to evaluation which views policy makers as an essential part of the audience for environmental and natural resource management research results.

### **What is the background to this study?**

LWA’s strategic plan for 2005-2010 was drawn up around three strategies; investing in strategic R&D targeted to Australia’s major national resource management issues, building on existing partnerships and brokering new partnerships, and translating existing and new knowledge into practice to achieve outcomes.

This outcome focus led to a linked focus on monitoring and evaluation (M&E), ensuring that project and program outcomes scaled up under each of the three corporate strategies and ultimately to the corporate outcome – demonstrating both accountability and purpose along the way.

In terms of monitoring the impact of their research, LWA took the view that ‘the return on investment in applied research is measured by the impact resulting from the adoption of the outputs of that research...applied research investors must be very attuned to the knowledge needs of intended end users of that research...knowledge and adoption strategies need to be ‘hard-wired’ into research program design and management...*the management of knowledge assets, long after the research phase has continued, is a growing responsibility of professional research investors*” (emphasis added).

Evaluating the impact of research was thus an integral part of LWA’s corporate strategy. Instead of seeing research projects as completed once the last payment was made to contractors, LWA decided to view them as live regardless of their age – and to evaluate their impact on their intended target audiences every few years. The corporate ‘communications’ strategy (3 per cent of total budget) morphed into a corporate ‘knowledge and adoption’ strategy which took 18 per cent of LWA’s total budget.

Evaluation of research was designed to answer five things (bearing in mind that research projects are still considered to be live well after the actual research work has finished):

1. What knowledge assets were generated? LWA explicitly recognised that different types of knowledge assets could be generated, including local, indigenous and strategic knowledge as well as the more ‘traditional’ scientific knowledge.
2. What does LWA know about the uptake and application of that knowledge among target groups?
3. What is known (or assumed to be known) about the impact of the application of that knowledge?
4. What impact has the adoption of this knowledge had on management and condition of Australia’s natural resources?
5. What is the return on investment of LWA and its partners across the whole research portfolio?

The objective of LWA’s knowledge and adoption strategy was to answer all five questions as thoroughly as possible, for all the target audiences. As well as the more traditional audiences often considered by NRM research outfits (e.g. extension workers), LWA explicitly included primary industries, indigenous communities and policy makers among its audiences and sought to identify the impact of their research on each one, knowing that each sought and used different types of knowledge asset from research.

Another important component of the LWA approach was to recognise that it is very rare for a single research project to generate all the knowledge assets needed for a particular end user. Instead, investors in this sort of public-goods research need to be able to assess how different research (even research carried out by different agencies) interacts over time to meet a given need.

**Have any particular theories or conceptual frameworks informed the development of the evaluation approach used?**

This approach is largely derived from practical need.

**What evaluation approach was used to assess policy impact?**

A targeted adoption model was developed separating out policy, on-ground practice (e.g. farm decisions) and planning, such as catchment or regional NRM plans & local government land-use planning. This detailed model forced LWA to consider M&E at all stages of the research planning and delivery.

Planning for evaluation of policy impact was done by assessing the knowledge-seeking behaviours of policy makers. Key issues identified were (see Bielak *et al*):

1. Targeting key ‘influencers’ at all levels, from Ministerial advisers to fast-track individuals and working out key times to interact with policy makers – rather than assuming that targets were likely to read summarised research reports, it was realised that (for example) senior policy people read emails on a Sunday evening at home, and prefer breakfast meetings for briefing (this list of key people is kept under review)

2. Working closely with policy people early in the process to scope research questions: including questions that arise from horizon scanning and foresight exercises
3. Developing specific knowledge management tools to make it easy to access information from their desktops (one-click tools)
4. Ensuring that research doesn't generate surprises – whilst not censoring uncomfortable findings, being sure that policy makers are briefed before any such findings hit the media – making for a more trusting relationship all round.

While this work is interesting and relevant to this project Land and Water Australia has now, however, been abolished and was planned to be fully wound-up by December 2009.

**Were any challenges highlighted in assessing the impact of the organisations work (and particularly of *environmental* research)?**

Several common issues were highlighted including: buy in of staff and policy-makers, costs, and attribution issues.

**Information source(s) for the case study**

Campbell, A. and Scholfield, N. (2006) *'The Getting of Knowledge: A guide to funding and managing applied research'*. Land and Water Australia, Australia.

LWA (2005) *'Knowledge and Adoption Strategy: managing information and knowledge for adoption outcomes'*. Land and Water Australia, Australia.

Bielak, A., Campbell, A., Pope, S., Schaefer, K., & Shaxson, L. (2008). 'From Science Communication to Knowledge Brokering: the Shift from 'Science Push' to 'Policy Pull''. In D. Cheng, M. Claessens, T. Gascoigne, J. Metcalfe, B. Schiele, & S. Shi (Eds.), *Communicating Science in Social Contexts: New models, new practices* (pp. 201-226). Dordrecht: Springer.