



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

## GUIDELINES AND SUPPORTING INFORMATION

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: guidelines and supporting information***

Version 1.2, 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 document was written by Ben Shaw and Sarah Bell of Policy Studies Institute, London with contributions from 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 the Guide and Guidelines or examples of the outputs they have resulted in 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

# How to evaluate the impact of environmental research on policy: guidelines and supporting information

## Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Planning a Research Impact Evaluation – 6 Questions</b>	<b>3</b>
1. Why are you interested in evaluating the impact of research on policy?	3
2. What will be evaluated?	4
3. When will the evaluation be undertaken?	6
4. Who will be involved in the impact evaluation?	7
5. Who will be using the evaluation outputs?	7
6. What resources are available for the evaluation?	7
<b>3. Five Evaluation Frameworks and Their Strengths and Weaknesses</b>	<b>9</b>
Method 1 - The Health Economics Research Group (HERG) Payback Model	10
Method 2 - The Research Impact Framework (RIF)	13
Methods 3 and 4 - The Irish Environmental Protection Agency (EPA) Framework and the Australian Research Quality Framework.	15
Method 5 - RAPID Outcome Assessment (ROA)	18
Summary of the Strengths and Weaknesses of the Five Evaluation Frameworks	20
Generic steps of evaluation	21
<b>4. Additional Evaluation Tools and Their Strengths and Weaknesses</b>	<b>23</b>
Complementary tools, indicators and metrics to draw upon to adapt existing impact evaluation frameworks	23
Five frequently used evaluation tools	30
<b>5. Evaluation Data Collection Methods and Their Strengths and Weaknesses</b>	<b>35</b>
<b>6. Generic Evaluation Issues</b>	<b>44</b>
<b>7. Worked Examples of Use of the Guide and Guidelines</b>	<b>46</b>
Worked Example 1: Developing a quick and light-touch impact evaluation framework for assessing research project and programme impacts.	46
Worked example 2: Developing a Strategic Impact Evaluation Framework	57
<b>8. References and Sources of Further Information</b>	<b>67</b>

# **How to evaluate the impact of environmental research on policy: guidelines and supporting information**

## **1. Introduction**

This document accompanies *How to evaluate the impact of environmental research on policy: a seven step Guide* (referred to as ‘the Guide’ through the rest of the document).

The Guide outlines a seven step approach for the development and implementation of a research impact evaluation.

This document is intended to provide further details to enable and inform the seven steps of planning and implementation of a research impact evaluation outlined in the Guide. It provides guidance for evaluating the uptake and impact of research on policy which will be of use to those using the Guide or those with general interest in the issues raised by the evaluation of research impact.

The final section of the Guidelines provides two worked examples of how the Guide is used. These illustrate how the Guide is used and give two outline evaluation frameworks suitable for use without use of the Guide.

The Guidelines can be read from cover to cover to gain insights into the issues surrounding planning and implementing a research impact evaluation, or specific sections can be referred to as required, for example, when details of the strengths and weaknesses of different data collection methods are required.

Not all the steps of the Guide require supporting information so the Guidelines are presented in the following main sections. The steps of the Guide the sections are most relevant to are:

Section 2 – Planning a research impact evaluation – Guide Step 1

Section 3 - Five evaluation frameworks and their strengths and weaknesses – Guide Step 3

Section 4 - Additional evaluation tools and their strengths and weaknesses – Guide Step 4

Section 5 – Evaluation data collection methods with case studies and their strengths and weaknesses – Guide Step 4

Section 6 - Generic evaluation issues – Guide Step 5

Section 7 - Worked examples of using the Guide and Guidelines – relevant to all steps.

More detailed information on much of this material is contained in the literature review and case studies compiled for the project. These contain full references for all the material cited and drawn upon and are available as separate documents.

## 2. Planning a Research Impact Evaluation – 6 Questions

*This section contains supporting material relevant to Guide Step 1: Identify Your Evaluation Objectives and Available Resources.*

Before an organisation can begin to consider how best to evaluate the impact of (their) research on policy, it is necessary to answer a number of context relevant questions, the responses to which will help inform the best research impact evaluation approach to adopt. Such questions include why, what, and when do they wish to evaluate, who should be involved in this process, and what resources are available for the evaluation. Six questions to answer are given in the Guide. These questions and the issues they raise in planning a research impact evaluation are discussed further below

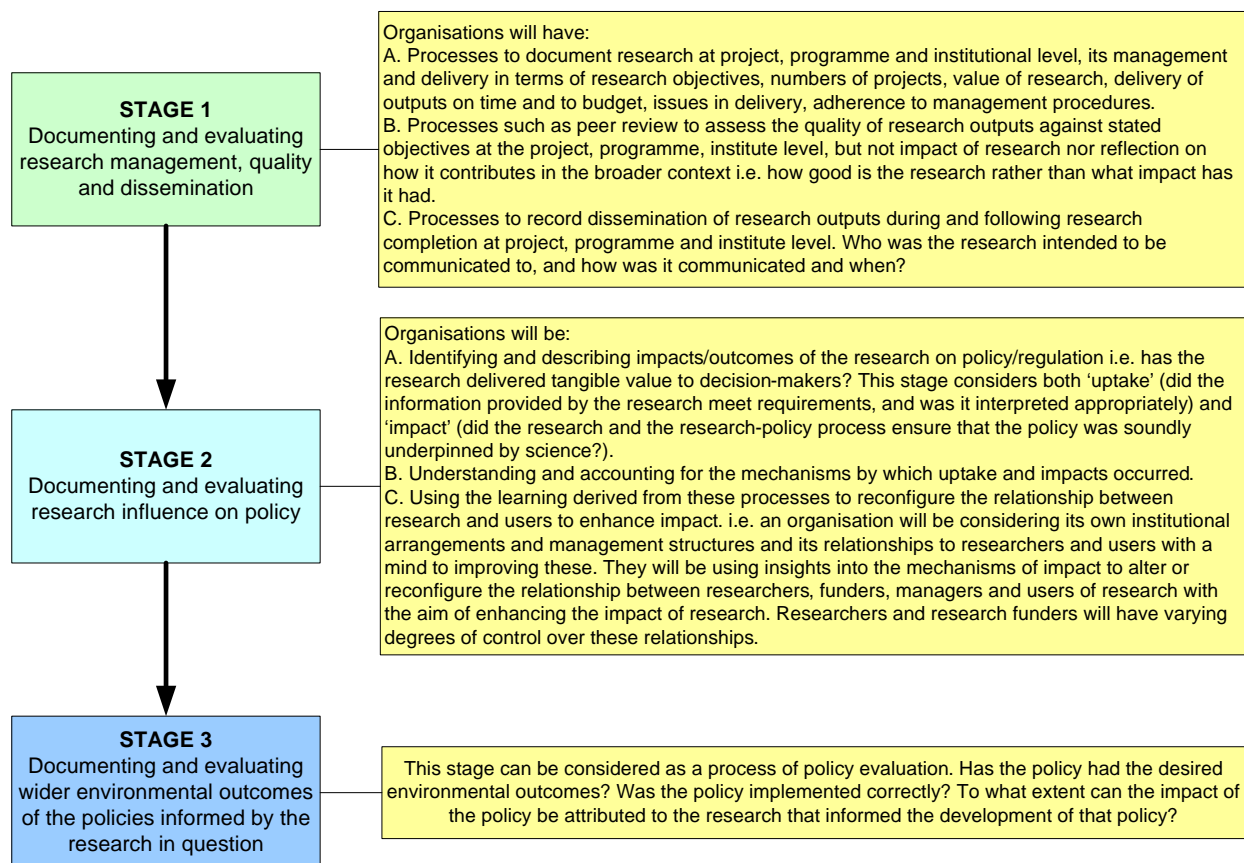
### 1. Why are you interested in evaluating the impact of research on policy?

Most research organisations currently evaluate research quality, and many evaluate the adherence of research management to procedures, but there appears to be an increasing aspiration to move beyond this towards an assessment of research outcomes and research impact (Lillis, 2000; Otronen, 2004). Multiple factors are driving this interest in impact evaluation, from a desire to increase the accountability of researchers to funding organisations and their use of public money; to the promotion of organisational research achievements; to learning within research institutions of how to broaden the ultimate reach and impact of their work.

In the current economic climate, impact evaluations are increasingly being perceived largely as a defensive process; they are needed not only to justify the value of investing in one particular research project or programme, but also to justify the value of *environmental* research over other types and, more broadly, to justify the value of spending public money on research itself, rather than on education, health etc.

Justification of investment is not, however, the sole reason why organisations are undertaking research impact evaluations. If undertaken properly, research impact evaluations can also constitute a powerful learning tool. The outcomes of the evaluation can be used to inform and enable organisations to restructure their existing research management processes and inject new ways of working into their institutions that will enable them to both monitor, and *enhance*, the impact of their research on policy processes i.e. evaluation is a tool for change, not just an end-product.

Figure 1 depicts three stages of organisational development with respect to research evaluation; it is the progression from Stage 1 to Stages 2 and 3 that many organisations are striving to achieve.



**Figure 1. Three stages of organisational evaluation activities**

## 2. What will be evaluated?

This question can be viewed from a number of different perspectives. Quite simply it may refer to the *type of research* whose impacts are being evaluated, for example basic or applied, or consumer-focused versus foresight-type research etc.

Alternatively, it may refer to the *level of research* which is being evaluated, for example, from an individual project, up through research programmes, institutes, to national/international research funding bodies. This will influence the scale and detail of the evaluation undertaken, potentially the level of resource that is available for conducting the evaluation, and the types of stakeholders that should be included in the evaluation as potential research users.

However, the most complex aspect of this question, and perhaps the most important in determining which impact evaluation approach to use, relates to the *type of research impact or outcome* to be evaluated, and the interest of the evaluator in understanding the mechanisms by which this impact occurred. The classification of impact adopted

will largely be shaped by the evaluator's conceptual understanding and interpretation of the relationship between policy and research, an issue which has been widely debated. Whilst some people suggest that research only has an impact on policy if there is a clear and direct link between research and policy outcomes (e.g. through the direct uptake of research recommendations by policy), others argue that this direct take-up rarely happens due to the myriad competing factors that influence the policy process. Instead it is suggested that a research organisation's greatest influence probably occurs by ***contributing high quality policy-relevant information to a pool of knowledge that policy-makers access when they need it and use as they see fit.*** It is most likely to be used to help them define the scope of problems and possible responses, rather than to dictate specific solutions.

Much of the research use in this latter model is not deliberate or direct and does not correspond to specific pieces of research; rather, bits of information seep into the mind, unrecognised and without citation. Information serves, cumulatively over time, a diffuse ***'enlightenment' function***, providing an understanding and interpretation of the data and the situation that is critical to the policy decision. As ideas from research become absorbed into conventional wisdom, they shape people's assumptions about how things work, about what needs to be done, and what solutions are likely to achieve desired ends. While this mechanism of research uptake has the potential to have a large impact, identifying it is a significant challenge for evaluation.

With this in mind, measuring the 'impact' of research solely by looking only at visible policy choices or policy outcomes would be misleading, or at least present an incomplete picture of impact, and should not be the only means of judging the impact of a research institute or programme.

Many different classifications of research impact on policy have therefore emerged, for example:

- A ***'continuum' categorisation of research use*** – e.g. Nutley *et al.* (2007) highlight a continuum of impacts from conceptual to instrumental, ranging from raising awareness, through shifts in knowledge, attitudes and culture, to actual changes in day-to-day practice and policies. This continuum is seen as a two-way process recognising that the use of evidence is likely to be iterative and interactive rather than necessarily straightforwardly linear. This builds on the widely referenced work of Carol Weiss (1979), which classified seven different approaches to research use, from the political and tactical to the enlightened.
- A ***'linear' categorisation of research use*** - e.g. research inputs lead to research outputs, which lead (possibly) to immediate outcomes, intermediate outcomes and then to ultimate (societal) outcomes. This, however, tends to suggest, either implicitly or explicitly, that there is a simple linear relationship between research and inputs, outputs and outcomes (i.e. that inputs lead sequentially and

logically to outputs and then to various stages of outcomes). In reality, the actual processes by which impacts occur tend to be more complex and not necessarily linear.

- ***A categorisation based on different aspects of impact*** – e.g. Raitzer and Ryan (2008) highlight three categories of impact: the *enhanced ‘diffusion’ of research* (this refers to citations and other measures of the degree to which there is awareness and transmission of research findings among different audiences); *greater ‘influence’ of research* (this refers to the degree to which perceptions, conceptions and/or actions of policy-makers have been changed or confirmed by research findings); and *‘tangible impacts’* (attempts to quantify the benefits of changes in policy attributable to research).

### **3. When will the evaluation be undertaken?**

In many ways, the decision about when to evaluate will depend on the types of impact or outcome of interest to the evaluator, the type of research being undertaken, the timing of existing evaluation procedures, or the organisational, management and political interest in evaluations. A common issue raised relates to the time lag between research completion and its wider policy impacts, which can range from almost immediate to very long term. In some areas research impact can take twenty or even fifty years to occur following completion although this longer timescale is not typical in the environmental sector. Once an impact has occurred, the persistence of its impact (i.e. how long the impact lasts) will affect the ease of its evaluation and the chance of the impact being identified or overlooked.

Impact evaluations are often carried out as research is completed or just after, which can be too early for any or all impacts to have emerged. On the other hand, with increasing time and distance from the research there is the problem of declining data availability; research impact becomes increasingly diffuse as it is traced from outputs to ultimate outcomes. It has therefore been suggested that impact evaluations should be undertaken at different points in time to capture the different ways in which research influences policy over time.

Setting clear research objectives and considering intended users from the outset of a research project or programme may assist efforts to evaluate the degree to which research has achieved its intended impacts. This is, however, less helpful in evaluating any unintended impacts of research, which tend to be particularly common in basic or ‘blue skies’ research. Time-lags to impact are likely to be greater and less predictable in basic compared to applied research, which has implications for both the choice of evaluation approach and its timing. On the whole, SKEP members are concerned with funding applied research which is broadly helpful in designing an evaluation approach.

#### **4. Who will be involved in the impact evaluation?**

This question will to some extent be determined by both the motivation for the evaluation and the impacts of interest. Where accountability is a key priority, it is suggested that independent evaluations should be carried out by external evaluators (Raitzer and Ryan, 2008), or at least that some form of independent impact verification is provided. The Finnish Environment Institute, for example, uses a ‘balanced scorecard’ approach in which both the research project leader and the policy supervisor in the Ministry of the Environment (the key user) independently score the research against a set of criteria and then discuss and explore any significant discrepancies between the scores allocated. (Further details of this approach are given below in Section 3, **Additional evaluation tools and Case Study 5 of the Case Studies report.**)

Broadly it is important that evaluations are carried out by those with sufficient knowledge and experience to understand the quality of the research and recognise its applicability to practical problems. For applied research in particular, this may require the input of both experts and potential users, which underlines the need to give careful consideration to who will supply inputs to the evaluation. Past experience tends to indicate that researchers, research managers and funders, steering groups, research users and broader stakeholders all have a role to play in outcome evaluation.

#### **5. Who will be using the evaluation outputs?**

Closely linked to Question 2 is the need to consider who will be *using* the outcomes of the impact evaluation. Will it be used internally, for example, to feed into the organisation’s management and strategy processes or to inform how best to restructure a particular team to better meet their research objectives? Is it being done for a board of funders to illustrate the benefits gained from their investment in the research and encourage further investment? The answers to these questions will help to determine which stakeholders should be involved in the evaluation, whether independent verification of impacts is required and so on.

#### **6. What resources are available for the evaluation?**

Understanding the resources available for the evaluation (e.g. the budget, the time allocation, the staff involvement and available staff expertise in evaluation) will help to establish any limits to the scope/scale of the evaluation and will highlight the extent to which the evaluation needs can be met. SKEP members may need to prioritise particular evaluation needs given the resources available.

Given the importance that the availability of resources has in shaping and constraining the development and delivery of evaluation framework remarkably little specific material was found in the literature review on this issue.

In absolute terms evaluation may seem to be a large cost with the literature studied suggesting figures of one-half to five percent of programme budgets being spent on programme evaluation and this range can be used as a 'rule of thumb' for what to spend. However, the literature also highlights that this expenditure can result in more effective research being commissioned and conducted which results in greater impact so evaluation is potentially a very cost-effective activity to invest resources in.

It is also highlighted that implanting an ongoing system of evaluation may have large start-up costs but once set up the marginal costs are low and the required data may be largely collected as part of project and programme management systems.

Allocation of evaluation resources at the outset of project and programme development is considered important in terms of ensuring capacity for and timely evaluation activity. It should be remembered that evaluation is not just an activity to be conducted following the completion of research activity. It can be run in parallel from the outset of a project or programme. Evaluations conducted in this way may not only be more efficient and effective overall but allow actions to be taken to address any issues revealed by the evaluation while the project/programmes are in operation. This obviously can improve the utilisation of research budgets and the level of impact achieved.

### **3. Five Evaluation Frameworks and Their Strengths and Weaknesses**

*This section contains supporting material relevant to Guide Step 3: Choose an Impact Evaluation Framework That Best Reflects the Criteria You Have Prioritised.*

One of the key findings of the literature review conducted for this project was that an evaluation methodology needs to be tailored to the circumstances of its use. The review also revealed a large number of approaches for structuring, interpreting and collecting evaluation data.

However, to develop an evaluation framework tailored to the needs and circumstances of the organisation in which it will be used requires expert knowledge of the options available and their suitability for different uses. The aim of Steps 2 and 3 in the Guide is to reduce the need for expert input to the choice of an evaluation framework. Step 2 asks users to identify the key criteria that their evaluation is intended to meet. Step 3 then establishes how five different evaluation methods perform against these criteria. This performance assessment allows the most appropriate evaluation framework to be selected.

This section of the Guidelines provides short descriptions of the five evaluation frameworks incorporated into Step 3 of the Guide. It also gives more detail on their strengths and weaknesses and is intended to give a feel for what is involved in using each of the frameworks. The five evaluation frameworks that have been identified that could be used or adapted to meet the evaluation needs and experience of SKEP Network members are:

- HERG Payback Model,
- Research Impact Framework,
- Irish Environmental Protection Agency (EPA) framework,
- The Australian Research Quality framework
- RAPID Outcome Assessment.

These five approaches have been chosen as they give a range of options which vary in scope, scale and detail of evaluation, have different emphasis on engaging with internal and external stakeholders or the presentation of quantitative or qualitative results. Except for one approach (the Australian Research Quality Framework), they have been used and adapted successfully in practice, display an element of flexibility in their use, and have the potential to be of value in addressing the needs of the SKEP Network members.

## Five Evaluation Frameworks

### Method 1 - The Health Economics Research Group (HERG) Payback Model

The Health Economics Research Group (HERG) Payback Model consists of two main conceptual elements. Firstly, it has five ‘payback’ categories that are the areas of interest in which impacts may have occurred due to research. Secondly, it has a payback model which indicates the relationship between the payback categories and the mechanisms by which paybacks occur, i.e. a model of the research process from pre-project to post-project and ultimate impacts. The payback model provides a standard framework for the capturing and analysis of research impacts and mechanisms of impact from multiple projects.

The approach was developed by the Health Services Research Group at Brunel University in order to examine the impact of health services research, but has recently been applied to other areas such as basic and biomedical research and social science research. This framework tends to employ a mixed methods approach, including both quantitative (e.g. bibliometric analysis including ‘grey’ literature) and qualitative methods (case studies, documentary analysis, interviews with key stakeholders, questionnaires).

The following summary is taken from Hanney *et al.*, 2000<sup>1</sup>.

The HERG model categorises the types of ‘payback’ (benefits from research) in five domains, most of which contain further subdivisions:

- 1) **Knowledge benefits**; the contribution to knowledge, be it new, confirmatory or just local evidence of something already widely acknowledged in world literature. This has traditionally been measured using peer review processes, but these can usefully be supplemented by bibliographic techniques and patent analysis.
- 2) **Benefits to future research and research use**; the place of these benefits in the model can be complex. Subdivisions may include: better targeting of future research; the development of research skills, personnel and overall research capacity; a critical capability to utilise appropriately existing research, including that from overseas; and, staff development/educational benefits.

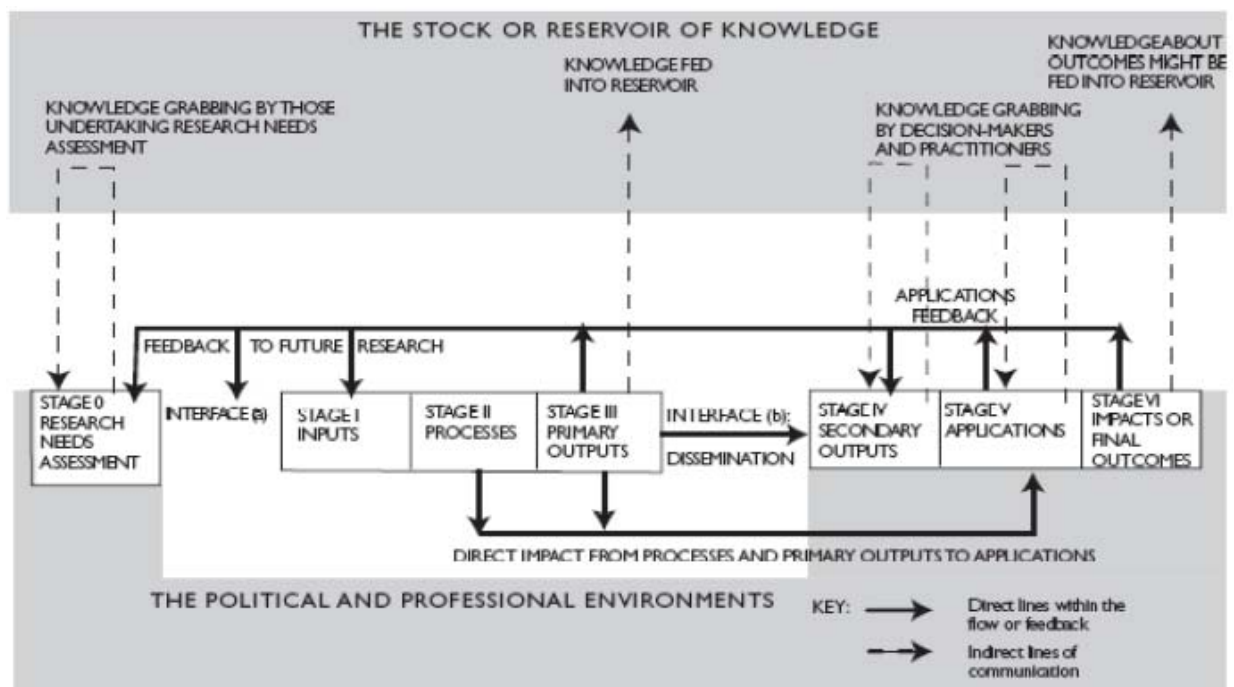
---

<sup>1</sup> Hanney S, Packwood T and Buxton M (2000), *Evaluating the Benefits from Health Research and Development Centres, A Categorization, a Model and Examples of Application*, Evaluation, Vol. 6, No. 2, 137-160, available from <http://evi.sagepub.com/cgi/content/abstract/6/2/137>. Similar material is available on an open access basis in: Hanney S, Gonzalez-Block M, Buxton M and Kogan M (2003) *The utilisation of health research in policy-making: concepts, examples and methods of assessment*, Health Research Policy and Systems 2003, 1:2, available from <http://www.health-policy-systems.com/content/1/1/2>

- 3) **Informing policy and product development;** these benefits include an improved information base and also evidence that decisions (e.g. about national or local policies, guidelines etc), were influenced by the improved information base. Other political benefits may include the use of research to: deflect criticism by showing the problem is being investigated; delay immediate decision-making; justify decisions taken for other reasons.
- 4) **Health (and health sector) benefits;** these benefits may be considered the ‘real’ paybacks from the perspective of those running health services and those who need to be convinced of the value of ‘diverting’ resources from provision of services to R&D. They could easily be adapted to explore *environmental* rather than *health* sector benefits.
- 5) **Broader economic benefits;** these include benefits arising from commercial exploitation of innovations arising from R&D.

It can be seen that the HERG approach considers a set of impacts of research which are much broader than just impact on policy.

Having created the multidimensional categorization of benefits, a model for assigning payback can be developed, as illustrated in Figure 2.



**Figure 2. The HERG Payback Model**

Ref: Hanney et al., 2000: p146

The first component of the model in Figure 2, **Stage 0 – Research Needs Assessment**, includes activities completed to shape and define the research prior to its commencement. The inclusion of **interface (a) – project specification, selection and commissioning** - highlights an interactive approach in which researchers assist customers in defining their needs, and in which expected dissemination and use of the research are negotiated from the beginning.

At **Stage I – Inputs**, it is important to consider financial inputs, experience of the researchers, the knowledge base to which they are contributing and the opportunity costs of their involvement. Interaction between customers and contractors is important at this stage, and even more so in **Stage II – processes**. Some outputs and paybacks may flow directly from the processes involved in undertaking the research before formal reporting/publication of findings because, for example, the behaviour of practitioners could change as a result of being involved.

At **Stage III – primary outputs**, knowledge benefits are prevalent, though these depend on the rigour of the science in many cases and can be complex. Better targeting of research, for example, requires a feedback loop to needs assessment of future research, and the training of researchers not only helps encourage the uptake of research findings but also feeds back as an input into future research. Some publications from Stage III will more appropriately be seen as part of **Interface (b) – dissemination**. Dissemination activities can be seen as part of the research flow but generally the findings are less likely to go directly to policy-makers, instead entering a much broader and looser stock or **reservoir of knowledge**. If decision-makers or practitioners are influenced by reading reports etc sent to them as part of the dissemination process, this may be considered as part of the flow. In contrast, where knowledge enters the reservoir of knowledge and is picked up for later application (be it in policy decisions or in research needs assessments etc), this is better considered as ‘knowledge grabbing’. The more research findings from one project mix with those from others and are later grabbed by policy-makers or practitioners, the harder it becomes to isolate and identify the benefits from any particular project.

Many of the paybacks occurring at **Stage IV – secondary outputs** can be classified as political and administrative. A decision in line with research could have been taken for a range of reasons, including: the substance of the findings influenced decision-makers; the findings enabled policy-makers to justify a decision they wanted to take for other reasons; the decision would have been made in the absence of the findings, but nevertheless they confirmed its legitimacy. Even where a policy decision is not totally consistent with research findings, they may still have played a major role in the policy deliberations.

Changes in the behaviour of practitioners/policy-makers have to occur before most of the final two categories of payback – health service and economic benefits – will be realised in **Stage V – applications**. At this stage, changes in behaviour have to be

recorded, and efforts made to identify how far any changes are a product of the research. Finally, at **Stage VI – impacts or final outcomes**, the major assessment of service and economic paybacks can be made using various measurements such as calculations of cost savings, satisfaction with the services provided etc. Even where observational measurements are possible, it remains important to demonstrate the link with the research project through attributable changes in use, against an assumed counter-factual.

## **Method 2 - The Research Impact Framework (RIF)**

The Research Impact Framework (RIF) is an easy and quick to use tool that uses descriptive impact categories to prompt researchers to systematically explore and describe the impact of their work. It was developed by the London School of Hygiene and Tropical Medicine in order to guide an analysis of the impact of a selection of their research projects.

The RIF is not itself intended to be evaluative in terms of prioritising impacts or proposing causal pathways. It seeks to: act as a standardised framework to help capture impacts across research topics and methods and facilitate comparison across time and cases; to guide researchers in planning research implementation and evaluation strategies; to assist researchers in looking at the broader influences and effects on and of their work in society; to promote research accountability in relation to the use of resources and the consequences of research; to help in the attribution of effects to health research given the range of other determinants of health and societal impacts; and to contribute to more extensive or specialised evaluations of research impact (Kuruvilla *et al.*, 2006).

The RIF is primarily a descriptive rather than explanatory approach, and is a four-dimensional framework providing an overview of potential (sometimes overlapping) research impact areas. The four narrative impact areas include:

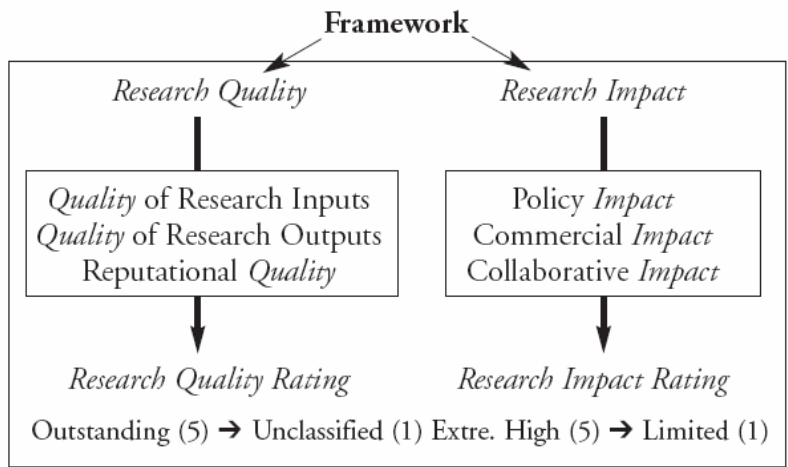
- 1) *Research-related impacts* e.g. research methods used, publications and papers, research networks, research management, dissemination and communication etc.
- 2) *Service impacts* e.g. evidence-based practice, quality of health care etc.
- 3) *Societal impacts* e.g. health literacy, health status, knowledge, attitudes and behaviour, social capital, sustainable development outcomes etc
- 4) *Policy impacts* e.g. level of policy-making influenced (local, regional, national etc), type of policy (e.g. practice, service, governance policies), the nature of policy impact (i.e. conceptual, instrumental etc), policy networks (the extent to which researchers are part of, or inform, policy networks), and political capital (the value of research evidence and researchers themselves in policy negotiations, in reaching high quality agreements and improvements to the policy-making process etc.).

Themes within each narrative impact area can be removed, added to, grouped, or modified as appropriate to the research being described. It is emphasised that researchers should be encouraged to think about the negative as well as the positive impacts of their work, such as stigma that could arise with publication of research findings related to a particular group or community.

**Methods 3 and 4 - The Irish Environmental Protection Agency (EPA) Framework and the Australian Research Quality Framework.**

The Irish EPA framework and Australian Research Quality Framework are closely related approaches. The Irish EPA framework is primarily an indicator-based approach, with indicators sitting under an umbrella of two main criteria: research quality and impact (see Figure 3). Each criterion is divided into three sub-criteria (each with several indicators) and data is collected on each of these indicators and used to rate both the quality and impact of the research in a simple numerical format. The three sub-criteria of research impact are:

1. *Commercial impacts*: the effects of the research on commercial activities such as patent applications or savings in resource use;
2. *Collaborative impacts*: this measures the linkages between the researchers and the policy community, commercial enterprises and other key stakeholders;
3. *Policy impacts*: this measures the contribution of the research to national and international policy development including work on advisory panels to industry or government agencies.



**Figure 3. Irish EPA Research Impact Framework**

*Ref: O’Leary et al (2008): p162*

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 written for each project and this allows the opportunity for other factors not captured by the indicators to be factored into the overall rating. This aspect builds in some element of flexibility to the fairly rigid framework.

Based on the data collected, both the quality and impact of the research is rated by an in-house panel of experts, using a five-point rating scale. This rating is used to produce two scores for the research; one reflecting quality and the other reflecting impact. Subsequent use of this approach in the Environmental Technologies Research Programme also 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). The Irish EPA approach flags up the potential for incorporating research **quality** into a theoretical framework.

The approach was largely derived from the Australian Research Quality Framework (RQF), which is more comprehensive but was abandoned prior to its implementation following Australia's 2007 election.

The key difference between the two approaches relates to the expert panel composition and the evidence used to inform its decision. The evidence portfolio required in the Australian RQF included: the context statement; four 'best' research outputs per researcher; a full list of research outputs; statements of impact that could be verified by qualified end-users of research; and other discipline-specific measures. The impact statement included: an evidence-based statement of claims for the Research Group against generic and panel-specific impact criteria, including verifiable indicators in support of these claims; up to four case studies that illustrate the Group's claims of impact; and, details of end-users who can be contacted by Assessment Panels to verify the Research Group's claims.

Both the Irish and Australian approaches use an expert panel to assess quality and impact based on the evidence portfolio. However, the Australian panel was proposed to be made up of 12 people, including a mixture of end users and internationally renowned peers i.e. is independent. In contrast, the Irish EPA one consists of two internal experts and one external Water Framework Directive expert (this is a practical response to the scale of evaluation and the resources available). Efforts are being made to engage a greater number of stakeholders in the Irish EPA evaluation process, primarily through consultation in the form of extensive meetings, emails and telephone communications with the research-project leaders, relevant EPA personnel and other key project stakeholders (Wemaere *et al.*, 2009).

While both approaches use context statements, the Australian RQF context statement is also accompanied by an in-depth impact statement, which includes verifiable, evidence-based claims against specific impact criteria, up to four case studies that illustrate examples of those claims, and details of end-users who can be contacted as referees to

verify the claims, again highlighting the more objective nature of the Australian approach.

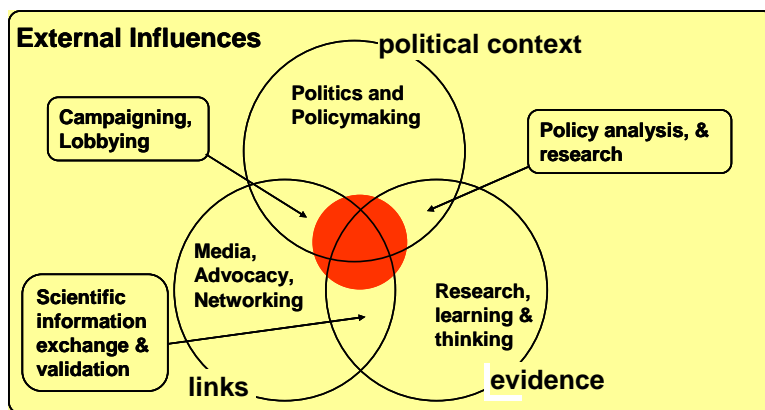
## Method 5 - RAPID Outcome Assessment (ROA)

RAPID Outcome Assessment (ROA) is a learning methodology to assess and map the contribution of a project's actions on a particular change in policy or the policy environment. It is a flexible and visual mapping tool which can be used alongside other evaluation tools and methods. The framework was developed by the CGIAR-affiliated International Livestock Research Institute (ILRI) and the Overseas Development Institute's (ODI) Research and Policy in Development (RAPID) programme, in response to the need to better understand the processes and mechanisms that lead to pro-poor decisions at the policy level. It combines elements from

- the International Development Research Centre (IDRC) Outcome Mapping approach (examines the progressive behaviour changes amongst the key actors that the project is directly influencing),
- episode studies (tracks back from a policy change to identify the factors that contributed to it),
- case study analysis (tracking forwards from research), and
- Most Significant Change (helps to identify and prioritise the key changes).

ROA is designed to assess the contribution of a project's actions and research on a particular change in policy or policy environment, both during and after a project is completed.

The method involves triangulation and integration of data collected using three different methods including classical case study (what was done, what happened), behavioural changes among key actors during the life of the policy process, and an examination of the factors contributing to that change interpreted within the ODI's *Context-Evidence-Links* framework (see Figure 4). This framework considers the process and influences from the perspectives of political context, role and use of evidence, role of linkages, and effect of the external environment.



**Figure 4. Context-Evidence-Links framework**

Ref: [www.odi.org.uk/RAPID/Tools/Toolkits/RAPID\\_Framework.html](http://www.odi.org.uk/RAPID/Tools/Toolkits/RAPID_Framework.html)

The Rapid Outcome Assessment method follows a three stage approach:

1. The first stage is a preparation stage, in which a documentary review and a series of informal conversations are carried out to develop an initial picture of the project's history and the intended changes. This will be guided by the Context-Evidence-Links framework, and includes case study (or project-based) information and episode study (or policy-change focused) information.
2. The second stage is a participatory workshop, in which participants (typically including the project team, key stakeholders and other external experts) identify the key policy change processes and causes of change in order to build a map of influences. The workshop covers the following:
  - Defining the policy environment at the start and end of the project/period.
  - Identifying key policy actors and 'boundary partners' that were considered influential in the process of ensuring impact. These may be individuals or institutions, and can be clustered into groups, for instance 'Civil society organisations', 'donors', 'the private sector' etc.
  - Describing the behaviour of the key actors/boundary partners that contributed to the change in the policy environment or policy, and the behaviour at the beginning of the project.
  - Establishing a time-line.
  - Mapping key behaviour changes for each of the key actors/boundary partners along the timeline, from the beginning to the end.
  - Mapping both the key changes in the project (including organisational changes, outputs and changes in behaviour), and the external influences including actions of strategic partners and other exogenous partners, during the same period.
  - Determining the level of impact/influence of (i) the project on the changes in behaviour of the key actors/partners, and (ii) of external influences on the changes in behaviour of the key actors/partners and the project i.e. exploring the links between the actors' behaviour changes and the identified events, including project activities, external influences and other actors' behaviour change.
3. The final stage involves a follow-up 'de-briefing' process allowing researchers to develop and refine stories of change – a 'timeline' - identifying key policy actors and events and their contribution to change. The timeline is used to identify key informants to follow up with in-depth interviews, which aim to help confirm the linkages and influences determined in the workshop and to assess the nature of the contributions to change.

## Summary of the Strengths and Weaknesses of the Five Evaluation Frameworks

Of the five approaches described, only the Irish EPA framework has been developed specifically for the evaluation of the impact of *environmental* research on policy, but the other frameworks have been successfully adapted for other contexts and appear to have significant potential to be used in an environmental context. Each framework has advantages and disadvantages and these are summarised in Table 1.

**Table 1. Key strengths and weaknesses of the five impact evaluation frameworks**

Framework	Advantages	Disadvantages
<b>HERG Payback Framework</b>	<ul style="list-style-type: none"> <li>• Common framework enables comparative analysis of multiple projects</li> <li>• Explores and explains impact</li> <li>• Mixed-method approach to counteract bias</li> <li>• Comprehensive and systematic method</li> <li>• Potential to score observed payback/benefits (although methodologies for this need further refining)</li> <li>• Applied and tested in a wide number of impact studies</li> </ul>	<ul style="list-style-type: none"> <li>• Costly and time-intensive (breadth of information)</li> <li>• Assumes all outputs have equal impact</li> <li>• Multi-dimensional categories risk double-counting and can make interviews repetitive</li> <li>• Does not fully explain/account for complex research-policy interface and non-linear impact of research on policy.</li> <li>• Developed and primarily applied in health care research</li> </ul>
<b>Research Impact Framework (RIF)</b>	<ul style="list-style-type: none"> <li>• Standardised structure that can be used to compare projects and programmes</li> <li>• Maps potential impacts (positive and negative) which can be used to clarify dimensions of use</li> <li>• Simple, quick and low cost approach</li> </ul>	<ul style="list-style-type: none"> <li>• Descriptive categories – analysis remains subjective</li> <li>• Lack of triangulation/cross checking</li> <li>• Developed and applied in health care research</li> </ul>
<b>RAPID Outcome Assessment Model</b>	<ul style="list-style-type: none"> <li>• Comprehensively maps cause and effect of key events</li> <li>• Assesses contribution of key actors</li> <li>• Mixed methods to counteract bias</li> <li>• Has been tried and tested in the international development field</li> </ul>	<ul style="list-style-type: none"> <li>• Costly and time-intensive</li> <li>• Does not capture the economic impacts of a programme</li> </ul>
<b>Australian Research Quality Framework (RQF)/ Irish EPA approach</b>	<ul style="list-style-type: none"> <li>• Assesses the quality and the impact of research</li> <li>• Can alter/add to criteria used as programme requirements evolve</li> <li>• Includes a context statement to account for impacts not covered by the criteria i.e. adds flexibility</li> <li>• Allows comparisons across research projects/programmes by including generic criteria, whilst also retaining panel-specific criteria to ensure rigour and relevance</li> <li>• Can be adapted for different needs (e.g. the Irish Environmental Protection Agency’s adaptation of the framework, which is less resource intensive but does not include independent individuals on the expert panel)</li> </ul>	<ul style="list-style-type: none"> <li>• Rating of impacts by expert panels can be subjective</li> <li>• Rating methodologies may need to be refined. For example, temptation to rate a project with a ‘very good’ impact on the basis of potential rather than actual impact</li> <li>• The indicators could become the sole focus of attention, whereas they are meant to act as a guide and aid to learning</li> <li>• Independence of the evaluation determined by composition of the expert evaluation panel</li> </ul>

--	--	--

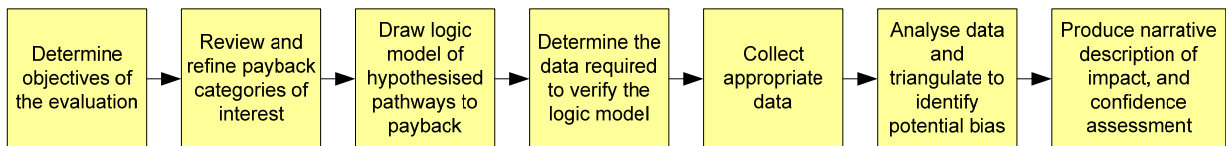
## Generic steps of evaluation

We have examined the five evaluation frameworks and broken them down into their steps of implementation from initiation to completion. The aim of this quick analysis was to gain insights into the processes used in each of the different frameworks and use this to inform thinking on how the different frameworks might be hybridised or combined to make them suitable for different evaluation circumstances. Breaking a framework down into smaller elements makes it easier to consider where its strengths and weaknesses lie and therefore how strengths might be transferred to other frameworks and weaknesses addressed. This analysis may be helpful in Step 4 in which the chosen evaluation framework is modified to better meet a user's circumstances and needs.

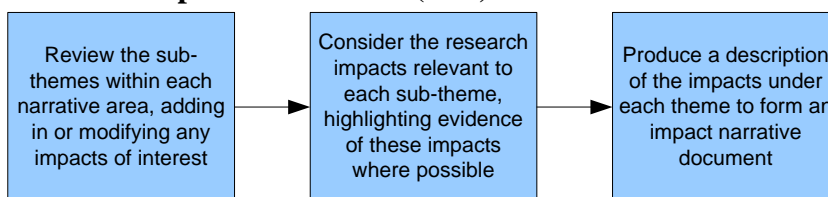
The elements of the five methods are illustrated in Figure 5. It can be seen that each method (very) broadly follows the six generic steps of implementation illustrated in Figure 6.

**Figure 5. Steps of Implementation of the Five Evaluation Frameworks**

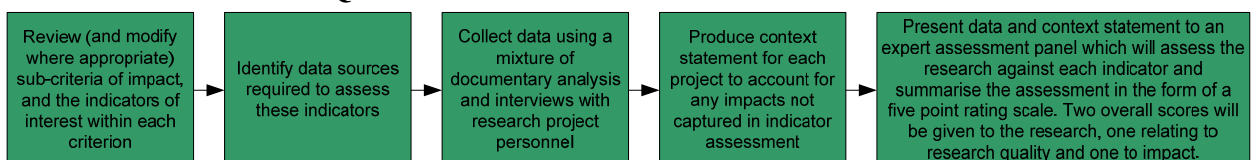
### HERG Payback Model



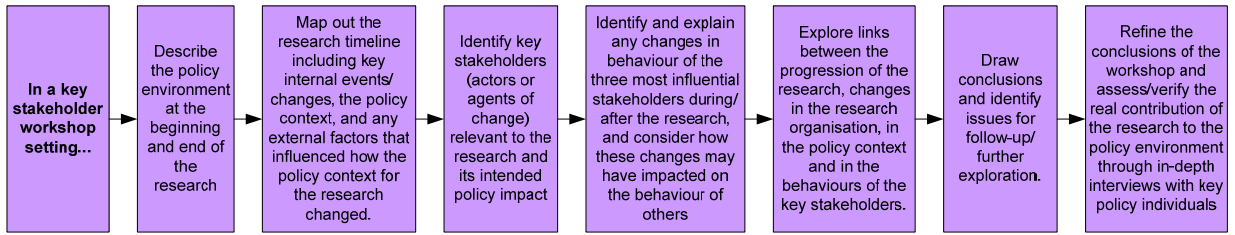
### Research Impact Framework (RIF)



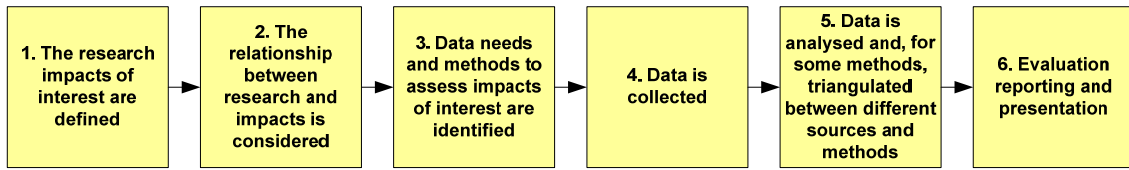
### Irish EPA/Australian RQF



## RAPID Outcome Assessment (ROA)



**Figure 6. Six Generic Steps of Impact Evaluation Implementation**



## **4. Additional Evaluation Tools and Their Strengths and Weaknesses**

*This section contains supporting material relevant to Guide Step 4: Refine Your Preferred Evaluation Framework*

Once the most appropriate evaluation framework has been selected in Step 3 using the assessment criteria selected in Step 2, it is likely that the framework will need to be refined to ensure it more closely aligns with the needs and constraints of the user. This may require the incorporation of complementary evaluation tools into the chosen framework to deliver elements of evaluation that are not yet considered or fully developed. For example, the Research Impact Framework (RIF) could be enhanced by using the double scoring approach of the Finnish balanced scorecard method, such that both the researchers and an external user rate the impact and quality separately. This would bring in a greater element of independence to the evaluation.

### **Complementary tools, indicators and metrics to draw upon to adapt existing impact evaluation frameworks**

Table 2 summarises 19 tools that have been identified by the literature review as being potentially useful to augment the five main frameworks, together with their strengths and weaknesses. They fall into three categories:

- a. Research impact evaluation tools.
- b. Indicators and metrics.
- c. Tools for research management and objective setting.

Following Table 2 further details are provided on the five evaluation tools that were most frequently identified in the literature review undertaken for this project.

**Table 2. Complementary tools, indicators and metrics to draw upon to adapt existing impact evaluation frameworks**

Evaluation method/tool	Strengths	Weaknesses
<p><b>Discrete research impact evaluation tools;</b> these include individual methods that can be incorporated into certain stages of the broader impact evaluation framework e.g. to explore potential areas of impact, to put an economic value on previously identified impacts, or to address the issue of attribution etc</p>		
<p><b>Impact pathway mapping/logic models</b> Involves the construction of hypothesised impact pathways, mapping out assumptions of where the information from research entered the policy-making system, the paths it took, and the decision-makers it influenced. These assumptions of transmission and uptake can then be explored or verified using multiple data collection methods and sources, including for example key informant interviews, and primary and secondary data sources and reports.</p>	<ul style="list-style-type: none"> <li>• Allows the pathways of impact to be hypothesised and data needs to explore and verify them to be specified</li> <li>• Can map routes towards impact, as well as actual impact.</li> <li>• Can highlight where research influence occurs but does not lead to tangible policy change (e.g. research findings used in policy debate but debate does not lead to a policy reform).</li> <li>• Can help highlight unexpected/unpredicted impacts.</li> </ul>	<ul style="list-style-type: none"> <li>• Risk that large domains of the logic model are missing or poorly defined</li> <li>• No formal statistical tests yet suitable for conducting significance tests of how the observed impacts match the hypothesised or intended impacts i.e. hard to assess causality/attribution</li> <li>• Can be labour/time-intensive</li> </ul>
<p><b>Episode studies</b> Uses a narrative to tell the history of events around a policy change. Uses a 4D framework (political context, use of evidence, links and relationships, external factors) to identify factors, key actors and events that have contributed to a policy change. Mixed methods used, including interviews and documentary analysis.</p>	<ul style="list-style-type: none"> <li>• Both direct and indirect observations</li> <li>• Assesses relative importance of influences</li> <li>• Process of working backwards can capture complex policy processes</li> </ul>	<ul style="list-style-type: none"> <li>• Over-emphasises political factors and under-emphasises the role of research.</li> <li>• Tracks backwards to consider factors influencing a particular policy change, so not as useful for organisations wishing to track forwards from their research.</li> <li>• Risk of actors ‘re-writing’ history</li> </ul>
<p><b>Simulation</b> This models mathematically the project/programme goals and constraints on meeting them to predict the impact of policy. Uses probability-based operations research modelling approach.</p>	<ul style="list-style-type: none"> <li>• Can reproduce variation and predict impact of multiple objectives</li> <li>• Can help choose between different options</li> <li>• Can be conducted <i>ex ante</i> or <i>ex post</i>.</li> </ul>	<ul style="list-style-type: none"> <li>• Complexity of model determines cost</li> <li>• Limited by the extent of understanding of the interaction between research and policy</li> <li>• Hypothetical, limited real world data</li> </ul>
<p><b>Economic analysis</b> For example, e.g. cost-benefit analysis, ‘willingness-to-pay’ techniques, partial equilibrium models. These approaches involve the statistical analysis of quantified inputs, outputs and outcomes. Often used to evaluate effectiveness of research programmes.</p>	<ul style="list-style-type: none"> <li>• Enables comparison of programmes on the basis of pre-established criteria</li> <li>• Discounted rates of return can account for time lags and weight factors according to</li> </ul>	<ul style="list-style-type: none"> <li>• Difficulties in quantifying outcomes, social welfare and externalities</li> <li>• Subjective interpretation of attribution and what to ‘cost’</li> </ul>

	<p>importance</p> <ul style="list-style-type: none"> <li>• Easily understood</li> <li>• Can help determine what benefits additional research funding will yield</li> </ul>	<ul style="list-style-type: none"> <li>• Resource-intensive method</li> <li>• Time separating research from economic benefit is often long</li> <li>• More applicable to applied than basic research; existing methods and data are sufficient to measure only a subset of important dimensions of the outcomes and impacts of fundamental science</li> <li>• Comparisons of cost-effectiveness of different projects can be difficult where very different types of output/outcome are produced.</li> </ul>
<p><b>Social analysis</b> Used to value the benefits of research according to quality of life/social rates of return relative to research inputs.</p>	<ul style="list-style-type: none"> <li>• Enables comparison of programmes on the basis of pre-established criteria</li> <li>• Discounted rates of return can account for time lags and weight factors according to importance</li> </ul>	<ul style="list-style-type: none"> <li>• Not all outcomes are quantifiable – requires subjective judgement</li> <li>• Measures impact of policy rather than impact on policy</li> </ul>
<p><b>Use of a counterfactual</b> Counterfactual scenarios can be generated to calculate what wider benefits of a particular policy would likely have been generated with all 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.</p>	<ul style="list-style-type: none"> <li>• Is regarded as a rigorous approach for trying to overcome the attribution issue</li> </ul>	<ul style="list-style-type: none"> <li>• Can be time/labour intensive</li> <li>• Difficulties in orienting interview and survey data around counterfactual identification</li> </ul>
<p><b>Network mapping and analysis</b> Analysis of the structure of relationships and the consequences for actors' decisions on actions. Mapping can identify multiple links (structure) and surveys/ interviews can explore how they are used and valued (agency). Can be examined from the perspective of a single actor or total network. (formalised survey, complemented by interviews, workshops).</p>	<ul style="list-style-type: none"> <li>• Reflects complex interactions of the realities of policy making</li> </ul>	<ul style="list-style-type: none"> <li>• Can identify linkages but cannot identify impacts/outcomes as a consequence</li> </ul>
<p><b>Positive utilisation narratives</b> A participatory method to identify stakeholder accounts of impact, and the use of secondary analysis to understand who contributed to change and how.</p>	<ul style="list-style-type: none"> <li>• Identifies unexpected change</li> <li>• Reflects organisational values</li> <li>• Quick and cost-effective method of evaluation</li> <li>• Self-evaluation and learning</li> </ul>	<ul style="list-style-type: none"> <li>• Anecdotal accounts and within-organisation evaluation prone to bias</li> <li>• Does not identify negative impacts or the non-use of research</li> </ul>

<p><b>Tracing post-research activity</b> Follows the impact of research according to channels of diffusion through networks and post-research activity of researchers.</p>	<ul style="list-style-type: none"> <li>• Broadens the scope of research ‘user’</li> <li>• Seeks to understand the dynamics of research flow, and interactions between key actors</li> </ul>	<ul style="list-style-type: none"> <li>• Difficulty in tracking activities</li> <li>• Anecdotal</li> </ul>
<p><b>Historical tracing</b> Tracing backwards from an outcome to identify contributing factors, using a range of (usually qualitative) data collection tools</p>	<ul style="list-style-type: none"> <li>• Policy-oriented approach to tracking events and processes</li> <li>• Can be used to explain ‘how’, ‘what’ and ‘why’</li> </ul>	<ul style="list-style-type: none"> <li>• Relies upon the quality of, and access to, existing documentation</li> <li>• Difficult to attribute causality</li> <li>• Does not account for indirect impacts, including dead-ends, spill-overs and synergistic effects.</li> </ul>
<p><b>Retrospective analysis</b> Related to case studies in that this approach also tries to reconstruct history; however it focuses on multiple scientific innovations/policy changes rather than just one. The goal is to identify linkages between innovations/policy change and particular types of antecedent events (usually either funding or research). Usually done by a panel of experts.</p>	<ul style="list-style-type: none"> <li>• Able to capture outcomes of research across many decades</li> <li>• Illustrates the linkages and factors necessary to advance research</li> <li>• Useful as a tool for distinguishing ‘policy churn’ (changes in the policy environment that may influence impacts) from economic impact i.e. some impacts last through changes of government or policies and others do not.</li> </ul>	<ul style="list-style-type: none"> <li>• Not as useful as a short-term tool because of the long interval between research and practical outcomes</li> <li>• Incorrect time interval may miss important factors or events</li> <li>• Dependent on human factors such as marketing of research and interest by professional audiences</li> <li>• Difficult to capture many contributing factors</li> </ul>
<p><b>Indicators and metrics approaches;</b> these can be incorporated as one aspect of a wider research impact evaluation e.g. to provide a numerical rating of impact against set criteria so that comparisons can be drawn across organisations or between the perceptions of different individuals involved in the evaluation process.</p>		
<p><b>Benchmarking</b> Uses baseline indicators against which to monitor progress towards objectives. Useful methodology for ongoing monitoring.</p>	<ul style="list-style-type: none"> <li>• Measures specific pre-defined outcomes</li> <li>• Provides a potential tool for comparison across programs and countries</li> </ul>	<ul style="list-style-type: none"> <li>• Objectives not necessarily measurable</li> <li>• Does not attribute causality</li> <li>• Policy has multiple-objectives; no single aggregate indicator appropriate</li> <li>• Can provide incentives to achieve</li> <li>• Very sensitive to the chosen indicators and may not take proper account of differing contextual conditions</li> </ul>
<p><b>Balanced Scorecard</b> A ten-step framework to evaluate research programmes/ organisations across four perspectives; employee learning, internal business, financial, and client relationships. Impact is just one of the categories evaluated.</p>	<ul style="list-style-type: none"> <li>• Focuses on direct attribution of a programme as it tracks immediate impacts</li> <li>• Performance indicators reflect specific goals and objectives</li> <li>• Learning, accountability and strategy</li> </ul>	<ul style="list-style-type: none"> <li>• Short term impact – less useful for assessing attribution of longer term outcomes</li> <li>• Useful for evaluating programme impacts, not necessarily impact on policy</li> </ul>

<p><b>Bibliometrics/citation analysis</b> Method for quantifying the impact of research by counting the number of outputs and citations, and analysing citation data.</p>	<ul style="list-style-type: none"> <li>• Suitable for repeated analyses and comparisons</li> <li>• Measures original research, not the programmes resulting from it</li> <li>• Identifies research deemed to be important by subsequent research</li> <li>• Traditionally considers academic citations, but extended analysis includes grey literature</li> <li>• Data easily available</li> <li>• Relatively inexpensive</li> <li>• Provides rapid results</li> <li>• Correlates with other evaluation methods</li> </ul>	<ul style="list-style-type: none"> <li>• Measures outputs not outcomes; citations by researchers only represent the first of many steps by which research is put to use by society.</li> <li>• Not comparable between disciplines</li> <li>• Quantity of output may not reflect quality; risks distorting importance of select publications</li> <li>• Limited role where publication is not the goal of research</li> <li>• Many uncertainties, and thus can only be used as a partial indicator</li> <li>• Time lag between publication and citation</li> <li>• Can be artificially influenced</li> </ul>
<p><b>Patents/new technologies</b> Where research may have patentable outcomes, this approach gathers data about the number and nature of patents.</p>	<ul style="list-style-type: none"> <li>• Useful to identify linkages between research and specific outcomes</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to compare between disciplines</li> <li>• Time lag after publication</li> <li>• Most commonly used to evaluate the wider impact of research on industry</li> </ul>
<p><b>Tools that focus on <i>ex ante</i> research management and objective setting;</b> these tools aim to enable organisations to <i>enhance</i> the impact of their research on policy, rather than specifically to <i>measure</i> it.</p>		
<p><b>Logical Framework Analysis</b> Logical framework analysis is a planning and management tool which aims to facilitate clear and specific thinking about what a policy programme aims to achieve, and how, and highlights factors which the achievement of these objectives depends on. The framework is 'logical' as it makes explicit the links between a programmes activities, outputs, purpose and goals and the assumptions that these links are based on i.e. what assumptions need to be true for 'activities' to deliver 'outputs' and for 'outputs' to deliver the 'purpose' of the programme etc. Data are presented in a matrix and mapped against indicators of achievement, means of verification and risks and assumptions.</p>	<ul style="list-style-type: none"> <li>• Evaluation is embedded in the planning process</li> </ul>	<ul style="list-style-type: none"> <li>• Does not integrate contextual analysis</li> <li>• Log frame does not represent reality</li> </ul>
<p><b>ROAMEF</b> This framework is recommended in the UK Treasury's Green Book on evaluation, and reflects the key stages of the policy process; Rationale, Objectives, Appraisal, Monitoring,</p>	<ul style="list-style-type: none"> <li>• Puts evaluation within a wider framework describing the policy process</li> </ul>	<ul style="list-style-type: none"> <li>• Does not specify methods for conducting evaluation</li> </ul>

Evaluation, and Feedback.		
---------------------------	--	--

<p><b>Results-Based Management and Accountability Framework</b>  Framework to monitor outcomes through the process of activity; results chain, resources, process, outputs, client reach, outcomes, and impact.</p>	<ul style="list-style-type: none"> <li>• Focuses on attribution by guidance through the ‘results chain’</li> <li>• Helps set target and track progress</li> </ul>	<ul style="list-style-type: none"> <li>• No prescribed methodology</li> <li>• Linear model – not reality</li> <li>• Useful for evaluating programme impacts, not necessarily impact on policy</li> </ul>
<p><b>Outcome mapping</b>  Participatory planning tool to identify and monitor progress towards intended outcomes (defined as changes of behaviour, relationships, activities or actions of boundary organisations) i.e. both ex-ante and ex-post elements to the approach.</p>	<ul style="list-style-type: none"> <li>• Assesses both process and outcome</li> <li>• Assesses contribution (not attribution)</li> <li>• Enables stakeholder dialogue, learning and consensus building, as well as accountability.</li> </ul>	<ul style="list-style-type: none"> <li>• Does not evaluate value for money</li> <li>• Observed behavioural change posits a link between cause and effect which may be impossible to demonstrate</li> <li>• Costly and time-intensive</li> <li>• Captures elements of policy implementation rather than focusing on ‘paper’ policy</li> </ul>

## **Five frequently used evaluation tools**

The following section provides details of five evaluation tools (as opposed to frameworks) that are frequently reported to have been used within the literature. These are

1. Economic approaches to evaluating the impact of policy-orientated research,
2. Logic models,
3. Indicator-based approaches,
4. Use of a counterfactual,
5. The balanced scorecard.

### ***1. Economic approaches to evaluating the impact of policy-orientated research***

Economic approaches are routinely used in the evaluation and appraisal of policy options to provide policy-makers with quantified estimates of the costs and benefits of different policy options. There is a wide literature on the possible approaches that can be used and their strengths and weaknesses.

Molas-Gallart and Tang (2007) report on a workshop organised by the UK's Economic and Social Research Council (ESRC) that discussed the application of economic approaches to the assessment of the impact of social and economic research. The paper provides useful summaries of four case studies intended to understand the impact of various ESRC research activities. It also addresses the question of how economic approaches could have built on these evaluations, particularly in the context of the policy expectation in the UK that research funders should justify the value gained from investment of public funds. The workshop considered economic approaches that could be applied to research impact evaluation including: willingness to pay; case study approaches; quantitative approaches such as econometrics and economic modelling; peer review; and international benchmarking. It concluded that all the approaches could in principle be used to demonstrate accountability of research funding. However, to learn about the impact of research on previous policy efforts, only case studies (as a basis for developing quantitative estimates of impact) and, to a limited extent, career tracking, (using an economic indicator such as salary differentials to track impact over time) could be used. Given these results are based on workshop discussions some caution should be exercised in adopting them uncritically.

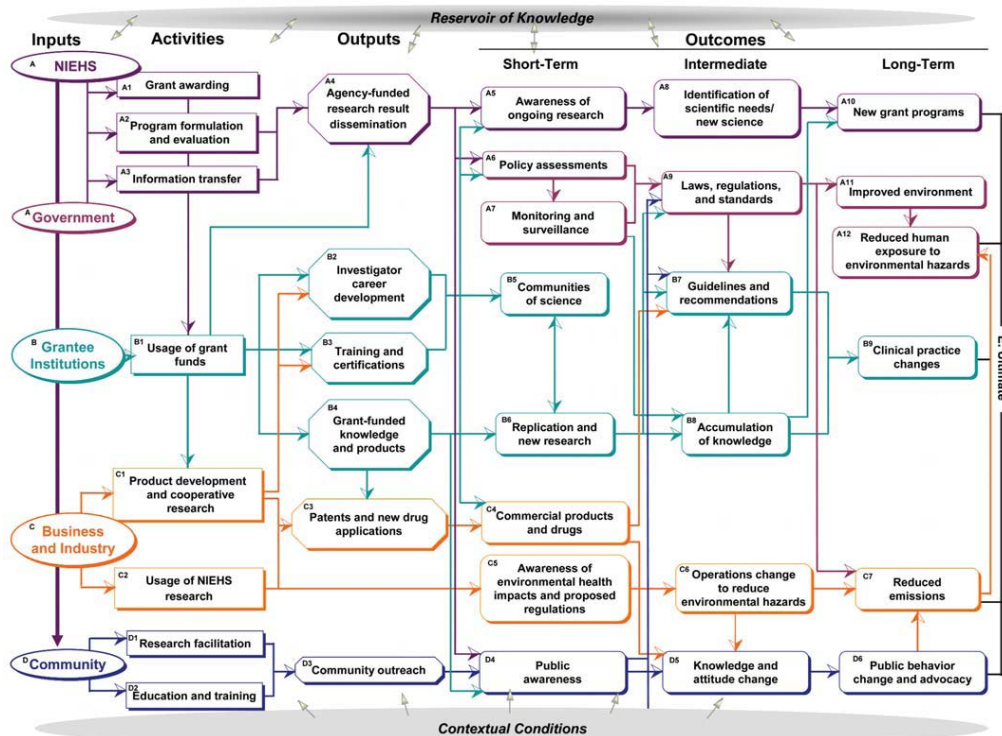
Ryan (2002) considers the impact of 'policy-orientated social science research' on policy. Impact is evaluated through a consideration of the degree to which research has advanced the speed of policy development and the economic benefits that can be attributed to this faster development. The approach combines qualitative interviews with the researchers and policy community and established that the research had resulted in policy decisions being taken earlier than they would have otherwise been taken (in the

Vietnamese rice commodity market). The economic impact of these earlier decisions was estimated using the partial equilibrium framework economic model to estimate the potential domestic and international welfare benefits of the research. While Vietnamese rice commodity markets may seem a rather obscure area, the approach illustrates how economic measures can be used to quantify impacts and how they can be linked with qualitative methods.

## 2. Logic models

A number of examples of logic models have been used and discussed in the literature, and the central element of the HERG model is its logic model. Engel-Cox *et al.* (2008) provide an example of how logic models, that is ‘*graphic depictions of the relationship between a program’s activities and its intended outcomes*’ - can be developed to assist in measuring the impact of environmental health research programmes on human health, the environment, and the economy, even when the impact may be indirect or diffuse.

The approach uses expert input and literature/documentary review on research impact assessment to develop a logic model that defines multiple components and linkages between (environmental health) research grant programmes and their outputs and outcomes in a variety of areas. The model delineates pathways, in which five types of institutional partners are included in the research process: the research funder, other government agencies, grantee institutions, business and industry, and community partners. Figure 7 illustrates the final logic model.



### **Figure 7. Example of logic model approach to research impact evaluation**

*(Source: Engel-Cox et al., 2008) Arrows represent linkages between the logic model components. Pathways are identified by letter and colour. Full details in original source.*

The approach of logic models has been generalised in the approach of the Logical Framework Analysis, which is a planning and management tool with four key dimensions: goals, purpose; outputs; and activities. Data are presented in a matrix and mapped against indicators of achievement, means of verification, and risks and assumptions. The two dimensions are linked by a set of logical prepositions and assumptions about the delivery of activities, outputs, purposes and goals.

### **3. Indicator-based approaches**

There are a number of indicator-based approaches to research impact evaluation described in the literature. The Canadian Academy of Health Sciences (CAHS, 2009) has developed an indicator based adaption of the HERG model. This has not yet been used, so its practicality is still to be tested. The approach proposed aims to define the impacts of Canadian health research, to learn how to improve the returns on investments in health research, to address gaps in the Canadian Institutes of Health Research framework, and to help clarify the ‘attribution issue of health research’. The key question asked: is there a ‘best way’ to evaluate the impacts of health research in Canada, and are there ‘best metrics’ that could be used to assess those impacts (or improve them?).

CAHS (2009) suggests a large number of possible indicators and metrics (numeric, descriptive and qualitative) covering a wide range of impacts. Interestingly from this project’s perspective is that CAHS are trying to create a framework that can be used in a non-prescriptive way by all the bodies who might want to evaluate the impact of their health research.

It highlights that health research generates a wide range of outputs with diverse outcomes. The uptake of that research is further influenced by a number of factors and demonstrating causal pathways that may lead from the research to impacts is consequently very difficult. It is therefore suggested that the problem of complexity creates the need for a standardised solution, involving the classification and mapping of the output-uptake-outcome relationships within an evaluation framework. The framework then allows identification of best indicators within various categories of impact. Using a standardised evaluation framework allows both comparison of evaluations and the identification of unexpected outcomes (as a framework can help to ensure that all possible outcomes are investigated).

#### ***4. Use of a counterfactual***

The Consultative Group on International Agricultural Research (CGIAR, 2008) uses a number of evaluation tools in their approach, drawing together impact pathways, economic approaches and the use of a counterfactual.

While each case is slightly different, the essence of CGIAR's approach is 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 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 policy-orientated research with the cost of the research to calculate the net present value, internal rate of return and/or the benefit cost ration, i.e. to identify the returns to the research investment.

#### ***5. The balanced scorecard***

The Finnish Environment Institute (SYKE), Finland's national centre for environmental research and development, conducts several layers of evaluation at different time-scales as described in Case Study 5 of the Case Studies report. A key feature of the annual evaluations is the 'balanced scorecard' process, in which the research project leader and the supervisor in the Finnish Environment Ministry discuss the success of the project including elements indicated in italics in the list below, which could be referred to as criteria of the uptake of them in policy making:

- The organisation of the project

- Budget issues
- Timetable issues
- How easy it was to agree things
- How well the project met its objectives
- The research methods
- *Networking aspects*
- *Its 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 scores the project leader is interviewed (recognising that more can often be learnt from the most extreme cases).

## 5. Evaluation Data Collection Methods and Their Strengths and Weaknesses

*This section contains supporting material relevant to Guide Step 4: Refine Your Preferred Evaluation Framework*

The Guide refers to the need to incorporate appropriate data collection methods into the evaluation framework and how different data collection methods can be used to overcome the limitations of particular methods. This section provides further information on the wide range of data collection options and their strengths and weaknesses.

Two case studies follow this information that illustrate how choosing the right combination of data collection methods can ensure important evaluation needs are met.

The framework developed in the Guide's previous steps will require data to complete the evaluation. Consideration will need to be given to where or who this data is sourced from and the methods by which it will be collected. There is a need to ensure that appropriate data for the chosen evaluation method(s) can be collected in terms of data availability and quality and the cost of collection.

Careful selection of data collection methods can help overcome certain limitations of the selected evaluation framework. This may be, for example, through helping to collect data on important evaluation criteria or impacts; or resource constraints could be addressed by using a less resource-intensive questionnaire/survey method to collect data on the HERG Payback categories rather than in-depth interviews.

Carefully selected data collection methods may also help to ensure that the more generic challenges of research impact evaluation are addressed e.g. issues of attribution and recall bias.

The literature review for this project highlighted a wide range of methods for gathering data that may be used in impact evaluations, which can be grouped into five broad categories:

1. *Qualitative methods*, including semi-structured interviews, documentary analysis, field visits and observations. These methods are used to generate rich descriptive and explanatory data sets that can be used to look at both utilisation pathways of research (i.e. mechanisms of impact) and the policy context in which utilisation is or is not taking place. However, they are often time and resource intensive, and it can be difficult to generalise their findings.
2. *Quantitative methods*, including surveys, bibliometrics and patent/new technology tracking. These methods are used to quantify the impact of research.

They are suitable for repeat analyses and comparisons, can be used to manage large amounts of data and can be relatively inexpensive. However, they are often difficult to use in the analysis of research impact specifically on policy. Survey response rates can be poor and bibliometric data usually focus on quantity, rather than quality, of research outputs rather than outcomes.

3. *Panels and peer review*. This is a relatively flexible and cost-effective approach to evaluation. Experts and peers bring status and credibility to the process and build ownership of the findings. These methods tend, however, to be used to enhance the quality of research rather than to assess policy impact, and panels are open to accusations of a lack of objectivity and are heavily reliant on the quality and experience of their membership.
4. *Workshops and focus groups*. This is an interactive, consensus building approach, but is not suitable for all topics (e.g. sensitive topics that generate polarised views) and can lack rigour and objectivity. It is, however, a relatively quick and low cost option for evaluation.
5. *Literature review*. This method is usually used as part of a mixed method study to scope a topic and place the impact evaluation in its wider context.

Further details of data collection methods and their strengths and weaknesses are included in Table 3. Some issues of particular importance to consider when choosing data collection methods are:

- **Speed and cost** – How quickly and at what cost can the data be acquired?
- **Availability** – Is primary or secondary data required? Are sources easily accessible? Is the source documentary or survey data, for example, and is it in the public domain, and/or are responses/participation from people required? Are the relevant people likely to be available or accessible and have the required time/willingness to participate? Are data available over the period of interest?
- **Reliability** – how reliable are the data collected likely to be? What biases might they introduce and how can they be overcome? Are there sampling issues that need to be addressed, e.g. are the right people being involved and can enough people be contacted to create a reliable picture(s) of research impact when users may be diverse and diffuse? Does an approach with broad coverage but less qualitative insight (e.g. bibliometrics or surveys) need to be balanced with an approach with narrower coverage but which reveals more depth, such as interviews?
- **Repeatability** – is it intended that the data be collected on a one off or a regular and ongoing basis?

Choosing the right combination of data collection methods can ensure important evaluation needs are met. Two examples, illustrating how this can be done are given in this section following Table 3.

**Table 3. Different approaches and methods for gathering evaluation data**  
(Details of the sources for this information are given in the Literature Review.)

Method	Description	Strengths	Weaknesses
<b>Semi-structured interviews</b>	A flexible interview around a framework of themes, with pre-identified key actors.	<ul style="list-style-type: none"> <li>• Flexible structure enables interviewees to respond in own terms, and interviewer to respond as part of a two way conversation</li> <li>• Framework ensures comparability of interviews</li> <li>• Seen as most appropriate when unravelling diverse layers and subtle nuances</li> <li>• Useful when there is a limited population of individuals with significant enough breadth and depth of knowledge to speak informatively about what happened and why (key informant interviews)</li> </ul>	<ul style="list-style-type: none"> <li>• Issues of bias and attribution</li> <li>• Training for interviewer is necessary (to avoid pre-empting answers or not probing enough)</li> <li>• Time-intensive (collection and analysis)</li> <li>• Interviewees may themselves be unaware of indirect influences of research</li> <li>• Time recall</li> <li>• Difficulties in locating appropriate key informants/research users</li> <li>• Often poor record-keeping by researchers of activities undertaken that may have encouraged research uptake (ephemeral connections), and by research users of specifically when/how they used the research.</li> </ul>
<b>Case study analysis</b>	An empirical approach that explores in-depth a project/programme, describing and explaining how and why developments of interest have occurred.	<ul style="list-style-type: none"> <li>• Can be descriptive and explanatory, and rich in detail</li> <li>• Can demonstrate pathways from research to impact</li> <li>• Potential to combine sources and methods (triangulation)</li> <li>• Explores context</li> <li>• Particularly useful in situations and contexts where the understanding of research impacts is poor, and there is only some knowledge on the causalities involved</li> </ul>	<ul style="list-style-type: none"> <li>• Difficulty in selecting cases</li> <li>• Issues of bias and attribution</li> <li>• Often over-emphasises importance of research – ‘supply perspective’</li> <li>• Time-intensive to ensure rigour</li> <li>• Difficult to apply a common framework across case studies. Must follow a standardised protocol in order to be comparable.</li> <li>• Often poor generalisability of findings</li> <li>• Highly dependent on the skill of the researcher</li> </ul>
<b>Documentary analysis</b>	Exploration and interpretation of existing documents. Can elicit quantitative or qualitative findings. Often used in conjunction with other methods.	<ul style="list-style-type: none"> <li>• Can be applied to a wide range of sources (including policy statements, technical reports, minutes, speeches)</li> </ul>	<ul style="list-style-type: none"> <li>• Relies upon the quality of existing records and access to these</li> <li>• No single methodology for analysis</li> </ul>

		<ul style="list-style-type: none"> <li>• Provides contextual understanding</li> <li>• Potentially cost-effective</li> </ul>	
<b>Peer/panel review</b>	Widely used advisory process of expert scrutiny of projects and programmes. Traditionally used to enhance or maintain quality of science, on the basis that experts in a particular field can reach a consensus.	<ul style="list-style-type: none"> <li>• ‘Experts’ confer status, credibility, and acceptability on findings</li> <li>• Can offer range of constructive feedback to guide process</li> <li>• Can be conducted at any time during the process of impact analysis</li> <li>• Relatively cost-effective</li> <li>• Flexible</li> <li>• Well understood</li> <li>• Useful for evaluating fundamental, long-term projects whose ultimate outcomes are unpredictable and not easily quantified</li> <li>• Can determine both quality and relevance of research</li> </ul>	<ul style="list-style-type: none"> <li>• Reflects upon rather than measures impact</li> <li>• Time-consuming method, particularly for the experts involved</li> <li>• Issues of objectivity and variability (including the quality of experts)</li> <li>• Impractical to evaluate a broad area (number of peers involved)</li> <li>• Criticised as eliciting acceptability rather than validity of findings</li> <li>• Qualitative findings only</li> <li>• Requires comprehensive information for reviewers</li> </ul>
<b>Surveys</b>	A pre-formatted series of questions asked of multiple actors, generating both quantitative and qualitative data.	<ul style="list-style-type: none"> <li>• Can identify outputs/outcomes associated with particular research</li> <li>• Cost-effective means of providing overview from range of actors</li> <li>• Wider range of stakeholders than possible with interviewing</li> <li>• Can identify greater breadth of impact</li> <li>• Can identify aspects to focus on in interviews</li> <li>• Useful for the triangulation of findings from other methods</li> </ul>	<ul style="list-style-type: none"> <li>• Relies upon access to respondents</li> <li>• Reflects the bias of those surveyed and those who respond</li> <li>• Unresponsive to unforeseen issues</li> <li>• May require follow-up interviews to fully understand the results</li> <li>• A written questionnaire approach can be too gross to capture nuances and subtleties surrounding influence</li> <li>• Can suffer from low response rates</li> </ul>
<b>Workshop, focus group</b>	An organised discussion with a group of individuals. The groups can involve a range of different stakeholders.	<ul style="list-style-type: none"> <li>• Can be conducted at any time during the process of impact analysis</li> <li>• Less expensive than surveys</li> <li>• Exploratory in-depth insights</li> <li>• Can reach a consensus that individual responses may not</li> </ul>	<ul style="list-style-type: none"> <li>• Risk of sample bias</li> <li>• Selective memory of participants</li> <li>• Unsuitable for competitive or sensitive topics</li> </ul>
<b>Literature review</b>	Synthesis of existing research relevant to the study. In the context of impact evaluation,	<ul style="list-style-type: none"> <li>• Useful initial research to define the scope of an impact study/ look at</li> </ul>	<ul style="list-style-type: none"> <li>• Depends upon ability to identify and access existing research</li> </ul>

	usually used with other methods.	evaluation methods <ul style="list-style-type: none"><li>• Cost-effective</li></ul>	
--	----------------------------------	---	--

<b>Field visit</b>	Primary research method where the research team visits in person the site of activity. Often used in the international development field to evaluate the broader impacts of a research programme. Can include observing meetings.	<ul style="list-style-type: none"> <li>• Can ensure information is up-to-date</li> <li>• Direct observation of activities</li> </ul>	<ul style="list-style-type: none"> <li>• Time-intensive and costly (requires planning and ex-post evaluation if it is to be beneficial)</li> </ul>
<b>User evaluations</b>	Participatory method for assessing stakeholder (either users of research or producers of research) satisfaction. May involve interviews, e-mail or telephone survey.	<ul style="list-style-type: none"> <li>• Looks at both research and research utilisation from a stakeholder perspective</li> <li>• In-depth understanding of utilisation processes</li> </ul>	<ul style="list-style-type: none"> <li>• Risk that stakeholders have a vested interest in expressing satisfaction</li> <li>• Can be both time-intensive and costly</li> </ul>
<b>Telephone interviews</b>	Usually semi-structured interviews, often used as preliminary means of identifying key stakeholders.	<ul style="list-style-type: none"> <li>• Can elicit open-ended information early on in research process</li> <li>• Enables greater sampling dispersion</li> <li>• Cost- and time-effective</li> </ul>	<ul style="list-style-type: none"> <li>• Less able to develop rapport with unknown interviewee (restricts line of questioning)</li> <li>• Interview length is limited</li> <li>• Cannot use visual prompts/sources</li> </ul>
<b>Impact Log</b>	A means of logging real-time direct impacts and uptake of research (i.e. from informal feedback, field observations).	<ul style="list-style-type: none"> <li>• Records impacts over time</li> <li>• Can be used to construct actor matrices</li> </ul>	<ul style="list-style-type: none"> <li>• Cannot be conducted retrospectively</li> <li>• Much of the information is anecdotal and subjective</li> <li>• Does not measure indirect impact</li> </ul>

Two examples of how data collection methods can be combined to improve an evaluation are given below. These are described in more detail in Case Studies 6 and 7 of the Case Studies report

### **Combining data collection methods example 1: New Zealand's Foundation for Research, Science and Technology (FRST)**

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.

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 of Ecosystem Research*

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's 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:

- 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.
- b. **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.

Information from the web-survey was intended, to some extent, to corroborate the more detailed but smaller number of interview reports

In summary, the data collection methods in this case study were chosen to allow a wide range of stakeholder input to the evaluation, but within the limited resources available, and included questions which enabled quantitative as well as descriptive analysis of impact.

### **Combining data collection methods example 2: The Environment Agency (England and Wales)**

This example concerns a set of five post project appraisals carried out by the Environment Agency (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.

The evaluations were carried out using a workshop format, rather than repeating past practice which was based on review of documentation and interviews. It was felt this would provide a better opportunity to draw out and explore issues. A two hour workshop was held for each of the five projects, with the aim of identifying lessons that could be learned to improve future project management practices, and in particular ensuring 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 they were 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.

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: attendees were given the opportunity to comment on it in draft to ensure it was a fair reflection of opinions expressed in the workshop.

In summary, this approach is rather like the Research Impact Framework approach, but by altering the data collection methods to incorporate a workshop format, it allowed new evaluation criteria, such as consensus building and opportunities for organisational learning, to be met.

## 6. Generic Evaluation Issues

*This section contains supporting material relevant to Guide Step 5: Check the Developed Evaluation Framework Meets Your Needs.*

The strengths and weaknesses of different evaluation frameworks and tools broadly fall into two categories. Firstly, there is a set of practical issues and constraints to be considered, such as those associated with the cost of an approach, the time required to conduct it and the availability of appropriate expertise. Secondly, there are more fundamental issues of evaluation practice and theory which relate to the limits and challenges of evaluation methods. These issues need to be considered if successful evaluations are to be conducted, and include: how to deal with attribution of policy impact to a research intervention; the effect of time-lags between research outputs and possible policy impacts; and the theories and understandings of the impact of research on knowledge and how knowledge is translated into policy.

Once the framework and associated data collection methods have been developed using Steps 1 to 5 of the Guide, the resulting approach should be checked to ensure that, firstly, it meets the needs and constraints identified in Step 1 of the process and, secondly, it deals with the generic challenges of conducting a research impact evaluation. This involves answering two questions:

1. Does the research impact evaluation framework developed meet the requirements identified in Step 1?
2. Does the framework developed deal with the main generic challenges associated with research impact evaluation?

Question 1 is addressed by returning to the questions in Step 1 and ensuring they have been addressed properly. Question 2 requires some consideration of the general issues associated with conducting evaluation generally and specifically in relation to research impact. Supporting material of relevance to these issues is summarised below, and they are explored in more detail in the literature review.

- **Time** – when is the evaluation best conducted to capture the impacts from research? There is an inevitable time lag between research intervention and policy impacts and, depending on the length of time passed, this can significantly increase the challenges of the identification and evaluation of impacts.
- **Subjectivity** – how does the approach deal with possible subjective data inputs or methods of analysis and/or triangulate results from different analyses either from within a particular framework or beyond it?
- **Complexity of the research/knowledge/policy interface** – how, if at all, does the framework account for the complexity of the interactions between research,

knowledge and policy and the broader context in which research is being conducted and policy is being developed? Is there an interest in understanding the impacts and/or the processes leading to them? (The relationship between research and impact was discussed in more detail above in Section 1. Planning a Research Impact Evaluation.)

- **Attribution** – does/can the framework attribute impacts to research interventions? If so how does it establish the mechanism(s) of impact? Or is the focus on establishing the ‘contribution’ of research to impacts? i.e. is the interest in establishing links between research interventions and impact, or is there an interest in identifying the causal mechanisms and significance of the research intervention relative to other potential influences and interventions?
- **Disciplinary basis** – does the approach rely on a single disciplinary basis and might this limit the scope of evaluation or confidence in results from it? Or does a multidisciplinary approach increase the complexity of evaluation? Can it consider and compare inputs and impacts in a range of ways, for example, in social, economic and environmental terms?
- **Comparability** - does the framework allow comparison of evaluations of different research activities – project, programmes, disciplines etc? Can evaluation outputs be compared with outputs from other frameworks now or in the future?
- **Unexpected or hidden impacts** – is the focus on defined outcomes of interest or identification of unknown outcomes? i.e. is the framework likely to reveal unexpected impacts (e.g. ones in non-target audiences) or hidden impacts (e.g. the easily overlooked impact of research that confirms existing policy positions)? Or is it assumed that impacts will occur in a specific area?

## 7. Worked Examples of Use of the Guide and Guidelines

### Two Complementary Approaches to Research Impact Evaluation

The project team recognises that not all the members of the SKEP Network interested in evaluation will necessarily want, or have time, to work through the Guide and Guidelines and supporting information. We have therefore used the process described in the Guide to develop two evaluation frameworks suited to the common research impact evaluation needs identified by the SKEP Network.

The need for these two complementary approaches was identified at the workshop held as part of this project in December 2009. The first was the requirement for a pragmatic ‘quick and dirty’ impact evaluation tool to be used where time and resources for research evaluation are limited. The second was the need for a more strategic impact evaluation tool which could be used to facilitate organisational learning and the better planning and management of research programmes and institutional structures in order to both monitor and enhance the impact of research outputs.

The outputs generated from each step of using the Guide are captured below to illustrate the process and assist those who may be going to use it. A summary of the resulting framework is given at the end of each worked example in boxes 1 and 2 and the steps of each framework are illustrated visually in figures 8 and 9.

### Worked Example 1: Developing a quick and light-touch impact evaluation framework for assessing research project and programme impacts.

#### Step 1 - Identify your evaluation objectives and available resources

This section describes the scenario in which we envisaged this evaluation framework would be used.

The primary motivation behind this evaluation is **to justify and demonstrate value** from the funds spent on a programme of probably largely commissioned research and, to some extent, **to promote overall organisational research achievements and impacts**. The need for evaluation is externally-driven and is therefore less concerned about ensuring organisational learning.

The key **focus is on *project-level* evaluations of the impact of *applied* research of direct policy-relevance**, and there may be multiple projects being evaluated within a larger overall programme. The projects are likely to have included impact-related objectives from their outset or commissioning, and performance against these should be explored in this impact evaluation. Impacts of interest are for the most part instrumental, i.e. they have a direct use, and primarily those impacts that occur within immediate and/or intermediate time-frames.

The evaluation will need to be undertaken and completed approximately six months following project completion, or less, and will largely be carried out internally rather than by external consultants. The framework could be integrated into the research management system and an annual process.

A research project manager (i.e. the manager of the research in the organisation conducting it) is likely to undertake the individual project evaluations with a contract manager (in the commissioning body) involved in the process. Research project and contract managers are likely to spend a maximum of half a day on the impact evaluation, but a programme evaluation coordinator will have longer to oversee the process. Key research users will be included in the evaluation but there will be limited engagement outside of the organisation.

The project-level evaluation outputs will be used by programme managers in the funding body and aggregated and synthesised to indicate overall programme-level impacts and benefits. This aggregated form should be suitable for presentation to funders and senior managers responsible for overseeing programmes and commissioning research.

Limited resources are available for the evaluation and it would be likely to be conducted in-house by researchers and research managers with limited evaluation experience.

## **Step 2 - Choose criteria to reflect key objectives of the evaluation**

Based on the above scenario and the list of criteria included in Step 2 of the Guide, the following criteria were rated as ‘essential’ in this impact evaluation scenario:

- Evaluates individual projects effectively,
- Can be conducted at low cost,
- Does not require a high level of internal evaluation expertise,
- Can be conducted quickly,
- Provides a summary impact score allowing comparison.

It is apparent that these are primarily concerned with resource availability for the evaluation and the need for comparability of evaluation outputs across different projects. An additional set of criteria were rated as ‘desirable’:

- Evaluates quality of the research
- Engages external stakeholders

- Describes and explains impacts. In this case the primary requirement is to supplement the impact score with a brief descriptive context statement, rather than an in-depth explanatory narrative.

### **Step 3 – Choose an evaluation framework that best reflects the criteria you have chosen**

Having scored the rankings given to each framework in Table 2 of the Guide according to the criteria prioritisations (above), the most appropriate evaluation framework in this case appeared to be the Research Impact Framework (RIF). This was on the basis of both its scores and a discussion of its pros and cons amongst the group using the Guide. However, RIF does not address *all* the essential and desirable criteria identified in Step 2, with key remaining weaknesses including:

- Failure to provide a summary impact score which would allow benchmarking and comparison,
- Failure to engage stakeholders sufficiently in the evaluation,
- Lack of consideration of quality of the research.

### **Step 4 – Refine your preferred evaluation framework**

Step 4 of the Guide gives instructions on how to refine the preferred framework identified in Step 3. The first step of this involves considering the individual elements of the framework.

In the case of RIF the key initial step of the evaluation framework is to define and amend the RIF research impacts of interest, based on a pre-existing list.

RIF was originally developed for use in the evaluation of the impact of health research. As such, it will be necessary to slightly modify the original impact themes and sub-themes included in the framework, and perhaps to add further ones where appropriate. A revised list of potential impact themes and sub-themes is given in Table 4 on the next page.

This table has a large number of impact themes and sub-themes. Users may want to narrow this down into a more manageable set and add themes or sub-themes to reflect their particular interests as required. The final agreed set of themes/sub-themes would be used to develop a short RIF impact questionnaire which will be used by the research manager and contract manager to rate the research against each of these impact impacts.

Where impact objectives are specified at the commissioning or inception stage of the project, the degree to which these have been met should be the focus of at least one question in the RIF questionnaire.

<b>Table 4. RIF Impact themes and sub-themes</b>	
<p><b>Research-related impacts</b></p> <ul style="list-style-type: none"> <li>• Type of knowledge/problem</li> <li>• Research methods used</li> <li>• Publications and papers</li> <li>• Products, patents and translatability potential</li> <li>• Research networks</li> <li>• Leadership and awards</li> <li>• Research management</li> <li>• Communication</li> </ul>	<p><b>Policy impacts</b></p> <ul style="list-style-type: none"> <li>• Level of policy-making</li> <li>• Type of policy</li> <li>• Nature of policy impact</li> <li>• Policy networks</li> <li>• Political capital</li> </ul>
<p><b>Service impacts</b></p> <ul style="list-style-type: none"> <li>• Type of services: environment and intersectoral</li> <li>• Evidence based-practice</li> <li>• Quality of environmental management and protection</li> <li>• Environmental information systems</li> <li>• Improvements in systems of environmental management</li> <li>• Cost effectiveness of environmental management and protection systems</li> </ul>	<p><b>Societal impacts</b></p> <ul style="list-style-type: none"> <li>• Knowledge, attitudes and behaviour</li> <li>• Environmental literacy</li> <li>• State of the environment</li> <li>• Environmental justice issues</li> <li>• Macro-economic/related to the economy</li> <li>• Social capital and empowerment</li> <li>• Sustainable development outcomes</li> </ul>
<p>Modified from Kuruvilla <i>et al.</i> (2006) which gives further details of each of the themes and sub-themes of impact and is available online at:  <a href="http://www.biomedcentral.com/content/pdf/1472-6963-6-134.pdf">http://www.biomedcentral.com/content/pdf/1472-6963-6-134.pdf</a></p>	

**Draw on additional evaluation tools/alternative data collection methods to improve the chosen framework**

The second approach indicated in Step 4 of the Guide to improving the preferred framework is to draw on additional evaluation tools and data collection methods.

To use the RIF approach, impact themes and sub-themes of interest are selected or if needed created. The researcher then notes any impacts resulting from the project in relation to each of the sub-themes. This information can then be used to create a short description of project impacts. The method is intentionally simple and was developed to be used by researchers who need to identify project impacts and present them concisely in, for example, project proposals or reports. As well as its simplicity the method is flexible as impact sub-themes can be developed or added to be relevant to the impacts of interest. However, the method can be susceptible to the subjectivity or

inconsistency of researcher opinions and data should be sort from alternative sources to verify the impacts identified by researchers.

From discussion amongst the group using the Guide three approaches were identified that could address the weaknesses of the RIF approach. These approaches were suggested on the basis of reading the case studies of past impact evaluations conducted by this project and documented separately, and the different options for data collection and the catalogue of additional evaluation tools, indicators and metrics in the Guide/Guidelines. The three additional approaches suggested are:

1. **Double scoring of impact sub-themes** (*adapted from the double scoring of the Finnish balanced scorecard approach*): both the research project manager and the contract manager should complete the short RIF questionnaire asking them to score the research in relation to each of the impact themes and sub themes. These scores may be added to provide a summary impact score, alongside the standard impact narrative produced by the RIF. (If resources allow a triple scoring approach could also be used that additionally asks an intended research user to complete the RIF questionnaire).
2. **Workshop** (*adapted from approach of the Environment Agency (England and Wales)*): a short workshop or evaluation meeting could be held to bring together a number of researchers working on the project, together with various potential users, to explore the impacts identified from the RIF questionnaires in more detail. The aim of the workshop would be to agree and verify the impacts identified by the research and contract managers with a wider group of researchers and users. This would enable greater stakeholder engagement and confidence in the impacts identified, particularly where scores diverged. Depending on budgetary constraints, and the number of projects to be evaluated in a programme, workshops could be held either for each project, for a set of projects on related themes, or to discuss impacts from the programme as whole. If resources are very limited just the research and contract manager could discuss and agree a final impact narrative and score.
3. **Research Quality**: commissioned research is likely to be subjected to various types of peer review, e.g. advisory panels, detailed review by individuals etc. The views of representatives of peer review groups should be obtained, and this could be done either as part of the workshop process in (2) above and/or separately by (telephone) interview.

### **Step 5 – Check the evaluation meets your needs**

This step is intended to check the outputs against the objectives from step 1 and general evaluation challenges. We have not recorded the discussion on these points as it would on repeat points made under the other steps. However, obviously it is important that

anybody using the Guide should conduct this step to ensure the framework developed is appropriate.

### **Step 6 – Implementation, detailed planning and data collection and analysis**

The data collection for this approach should be fairly straightforward with the RIF questionnaire possibly being integrated into the requirements of research project management systems and perhaps being conducted on an annual basis for all live and recently completed projects. Integrating the approach into project management systems would have the advantage that information on the status of the project – e.g. project name, objectives, cost of project, status (in progress/date of completion, etc, would exist alongside the impact information.

Online form-building/survey websites could be used to improve the efficiency of data collection, although sensitivities around data security should be considered.

A timeline for data collection will need to be determined and the overall evaluation coordinator will need to monitor response rates from research and project managers (and users if surveyed). A small incentive for return or sanction for non-return may be necessary to improve response rates.

The workshop(s) would be based on inputs from RIF questionnaires and from peer-review processes on project quality. It/they would need to be held following completion of the questionnaire phase allowing sufficient time for collation and initial analysis of responses. The main objective of the workshop(s) would be to act as a means of testing the reliability of the findings from the RIF questionnaires and identifying further impacts for reporting, rather than exploring mechanisms of impact in detail.

### **Step 7 - Evaluation Reporting, dissemination and reporting**

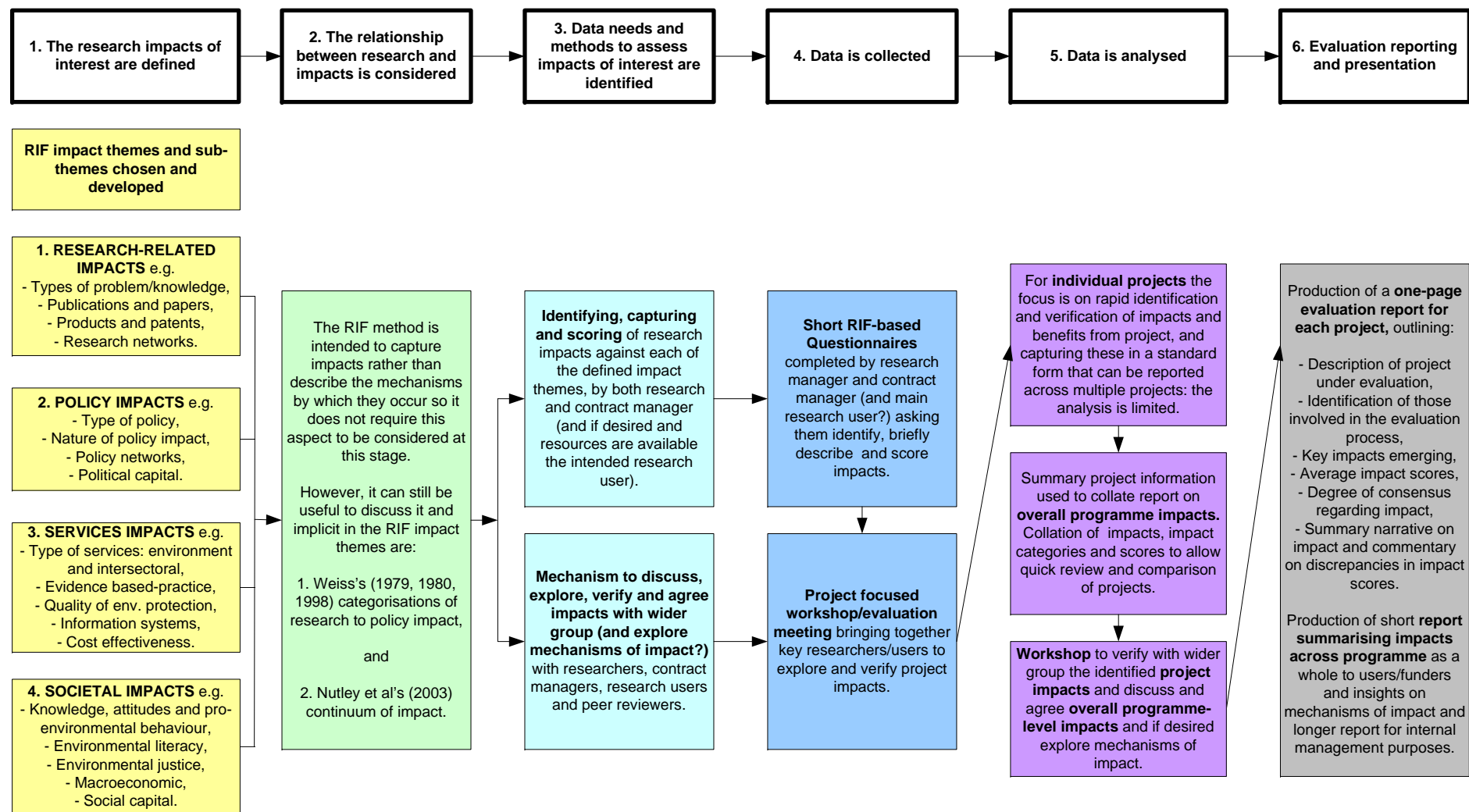
The outcomes of the project level impact evaluations could be reported under the following headings in a standardised short summary format (one page) document:

- Description of project under evaluation including intended objectives if set,
- Identification of those involved in the evaluation process,
- Impacts identified,
- Average impact scores,
- Degree of consensus and confidence in identified impact and attribution of impact,
- Summary narrative on project impacts and explanation of any discrepancies.

These reports would be submitted by the contract manager to the programme evaluation coordinator, who will synthesise the evaluation outputs across all projects. As a large number of project reports may possibly need to be compared they should be short and/or include a one page summary and be in a standard format.

A report summarising key project impacts across the programme and other relevant information should be drafted for discussion and development at a workshop to discuss overall programme impacts and benefits. A final report would be drafted incorporating the insights from this workshop and this, and/or a summary of it, would be submitted to senior programme managers and funders.

**Figure 8. Adaptation of the Research Impact Framework (RIF) to produce a quick and light-touch impact evaluation framework for assessing research project and programme impacts.**



**Box 1. Summary - a quick and light-touch impact framework for evaluating research project and programme impacts**

***When to use this approach:***

This approach can be used when resources (including funds, time and expertise) are limited and it is suitable for externally-driven, project-related impact evaluations which are to be carried out in-house by researchers and research managers with limited evaluation experience. The primary motivation behind this evaluation is to justify and demonstrate value from the public funds spent on a programme of (largely) commissioned research. The focus is primarily on the instrumental impacts of applied research, particularly those occurring in the immediate or intermediate term.

***Key evaluation needs:***

‘Essential’ criteria considered when developing this impact evaluation approach included the ability to: evaluate projects effectively; be conducted quickly at low cost without the need for significant internal evaluation expertise; and to provide a summary impact score. ‘Desirable’ criteria included the ability to: consider the quality of research; engage external stakeholders; and to supplement the impact score with a brief description of project impacts.

***The evaluation approach:***

An adapted version of the Research Impact Framework (RIF) (see Section 3 for details) best meets the needs of this type of evaluation. Adaptations to the basic RIF seek to enable a summary impact score to be produced, allowing comparison of multiple projects, and to encourage greater user-engagement in the evaluation process than in the original RIF. This approach therefore consists of:

1. Refining the RIF impact themes and sub-themes of interest to focus on environment-related impacts rather than health: see Table 4 above.
2. Development of a short impact questionnaire, based on the RIF impact themes of interest, to identify, capture and score project impacts.
3. Double scoring of impacts in each RIF sub-theme: both the research project manager and the research contract manager will be asked to complete the impact questionnaire. Identified impacts will be described briefly (a sentence or two per impact) and scored in each of the defined impact sub-themes to give an overall project score and an indication of the types of impact resulting from the project. Impact questionnaires could also be sent to the intended research user(s) to allow triple scoring of impacts to increase confidence in results if desired and resources allow.
4. A project impact summary report is produced (approximately one page in a standard format) by the contract manager bringing together the material from the returned impact questionnaires.
5. Short workshop/evaluation meetings with researcher team, contract manager and research users held to discuss summary report, reconcile any divergent views on impacts reported and scored and agree a final project summary report.



6. The evaluation manager collates summary impact reports from all projects evaluated and prepares a draft report on overall programme impacts and benefits, e.g. summary of project impacts; ranking of projects and key categories of impact, discussion of broader project and programme benefits and confidence in them; and any lessons on mechanisms of impact.
7. Workshop to bring researchers, contract managers and users from across the programme together to discuss and verify project impacts identified by the impact questionnaires and discuss overall programme impact and any other relevant issues.
8. Production of final evaluation report.

***Evaluation reporting:***

Upon completion of the data collection, a short one-page impact summary will be produced for each project evaluated, structured around standard headings, e.g. description of project under evaluation; identification of those involved in the evaluation process; impacts identified; average impact scores; degree of consensus regarding impact and attribution of impact; summary narrative on impacts and commentary on any discrepancies about impact.

This report will be submitted by the research project manager to the programme evaluation lead who will synthesise the project evaluation outcomes in a final report to senior programme managers and funders.

## **Worked example 2: Developing a Strategic Impact Evaluation Framework**

### **Step 1 - Identify your evaluation objectives and available resources**

The primary motivation behind this evaluation scenario is **to facilitate organisational learning about how to enhance the impact of research projects**, and the effectiveness of existing impact-related activities of a research organisation. It is likely that the motivation for the evaluation will be internally driven.

The key focus is on *programme*-level evaluations of the impact of both applied and basic research. The projects within the programme of research will not necessarily include impact-related objectives from the outset, and research users may not always be obvious. The evaluation seeks to explore the ‘enlightenment’ function of the research programmes as well as more tangible, instrumental impacts.

The impact evaluation is likely to form part of a wider strategic process. It may pick up on immediate or intermediate impacts of recently completed projects within the programme, but also on longer term impacts that are emerging from older research projects. Approximately six months are available for the completion of the impact evaluation, and it is expected to engage a much wider array of stakeholders and users, particularly those that are external to the research programme or environment agency or ministry commissioning or funding the research.

The programme evaluation outputs will be used by researchers, contract managers, programme managers and planners looking to enhance the impact of their research, and those involved in higher-level management, and may also contribute to the existing knowledge base about mechanisms for enhancing the impact of research on policy and/or industry etc.

A relatively large amount of resource is available for this impact evaluation, including full-time employment of a skilled evaluation lead for three-to-six months, plus additional time to account for the time commitment of interviewees and external users. The literature review conducted for this project suggested one-half to five percent of programme funding should be allocated to evaluation but given the relatively detailed evaluation scenario proposed the cost of the approach is likely to be towards the higher end of this range.

### **Step 2 - Choose criteria to reflect key objectives of the evaluation**

Based on the list of criteria included in Step 2 of the Guide, the following criteria are rated as ‘essential’ in this approach:

- Evaluates programmes effectively,

- Provides opportunities for internal consensus building and organisational learning,
- Engages external stakeholders,
- Describes and explains impacts,
- Produces results with high confidence and rigour.

It is apparent that these are primarily concerned with rigour, breadth, depth and confidence in the evaluation outcomes, and with gaining an in-depth understanding of how and why different impacts were realised. An additional set of criteria were rated as ‘desirable’:

- Considers short-term to long-term impacts,
- Is independent,
- Considers quality of research,
- Can be conducted at relatively low cost and relatively quickly.

### **Step 3 – Choose an evaluation framework that best reflects the criteria you have chosen**

Having weighted the rankings given to each framework in Table 2 of the Guide according to the criteria prioritisations (above), the most appropriate approach identified was the HERG Payback Model. However, the HERG Payback Model does not address *all* the essential and desirable criteria, with key remaining weaknesses including:

- The original HERG Payback Model is resource intensive and therefore a scaled-down approach will need to be developed.
- There are relatively limited opportunities for consensus building and organisational learning.

### **Step 4 – Refine your preferred evaluation framework**

Step 4 of the Guide gives instructions on how to refine the preferred framework identified in Step 3. The first step of this involves considering the individual elements of the framework.

The HERG Payback Model consists of two main conceptual elements. Firstly, it has ‘payback’ categories that are the areas of interest in which impacts may have occurred due to research. Secondly, it has a payback model which indicates the relationship between the payback categories and the mechanisms by which paybacks occur, i.e. a model of the research process from pre-project to post project and ultimate impacts. The payback model provides a standard narrative framework for the analysis of impacts and mechanisms of impact from multiple projects. The payback categories and payback model can be developed as part of the evaluation and they inform the data collection

approaches, for example, case studies, interviews, document or literature review and bibliometrics.

The HERG Payback Model was originally developed for the evaluation of health research impact. As such, it will be necessary to modify the original impact ‘payback’ categories included in the framework, to be relevant to the research and interests of the SKEP Network. This could result in the following payback categories:

1. Knowledge benefits,
2. Benefits to future research and research use,
3. Informing (environmental) policy,
4. Environmental benefits,
5. Broader economic and societal benefits.

Further categories, or sub-categories could be added, for example, on the basis of the RIF impact and sub-impact themes (see Table 4 above and the discussion of RIF impact themes and sub-themes in the quick and light-touch impact evaluation framework above.

### **Drawing on additional evaluation tools/alternative data collection methods to improve the chosen framework**

The second approach suggested in the Guide to improving the preferred framework is to draw on additional evaluation tools and data collection methods. The HERG Payback Model is a flexible approach which can use any appropriate data collection methods to inform the payback categories and model. So the focus here is on choice of tools rather than adding extra ones.

Based on discussions held amongst the group using the Guide three evaluation elements were identified to address the identified weaknesses of the HERG Payback Model in this evaluation context. These elements were chosen on the basis of the information included within the case studies of past impact evaluations, and from the Guide and Guideline’s options for data collection and the catalogue of additional tools, indicators and metrics. The elements are:

1. **Light-touch ‘payback’ questionnaire** – to obtain breadth of input from potential research users in a relatively low cost manner. This will involve asking researchers, contract managers and research users to complete questionnaires, possibly web-based ones, on research impact and mechanisms of impact. The aim will be to capture the detail, range and type of impacts across the programme and identify projects or issues for more detailed exploration in case studies and a workshop.

**2. Project Case Studies** – these are suggested to obtain detailed insights into research impacts and the mechanisms by which they occur, captured and reported in the framework of payback model. There is a need to conduct enough case studies to inform the key elements of interest in the payback model. Depending on the availability of resources the case studies could be based on a range of methods including documentary review, literature review, interviews and/or bibliometrics.

**3. Workshop** – with researchers, contract managers and a range of research users to:

- Present payback categories and payback model,
- Verify and explore impacts and findings of the questionnaires and case studies and identify any further impacts,
- Explore pathways and mechanisms of impacts proposed following the analysis of case studies and questionnaires and how to enhance these in future,
- Engage wider group of researchers and users in process.

### **Step 5 – Check the evaluation meets your needs**

This step is intended to check the outputs against the objectives from step 1 and general evaluation challenges. We have not recorded the discussion on these points as it would on repeat points made under the other steps. However, obviously it is important that anybody using the Guide should conduct this step to ensure the framework developed is appropriate.

### **Step 6 – Implementation, detailed planning and data collection and analysis**

While the data collection methods appear to be relatively straightforward, the challenges and resources required to collect and analyse data of this sort should not be underestimated. Identifying and contacting potential research users is time consuming. This is particularly the case when the research being evaluated is conceptual rather than instrumental, i.e. the focus is typically on evaluation questions such as, ‘how has this research changed what you think?’, rather than, ‘how has this research changed what you do?’. Additionally, the users of conceptual research may not be obvious and/or may be diffuse.

The evaluation will need to be carefully phased as the initial questionnaire will inform the case studies and both of these will need to have been completed and analysed, at least on an interim basis, before the workshop.

The analysis will require different sources to be compared, contrasted and, if necessary, conflicting findings reconciled or chosen between on the basis of judgement. The qualitative data generated are likely to generate conflicting and sometimes inclusive findings. This will require judgement on the part of the evaluator and for ideas and mechanisms to be proposed for testing with wider stakeholders at the workshop. Indicating the confidence in different findings from the process may be a useful way of presenting the results.

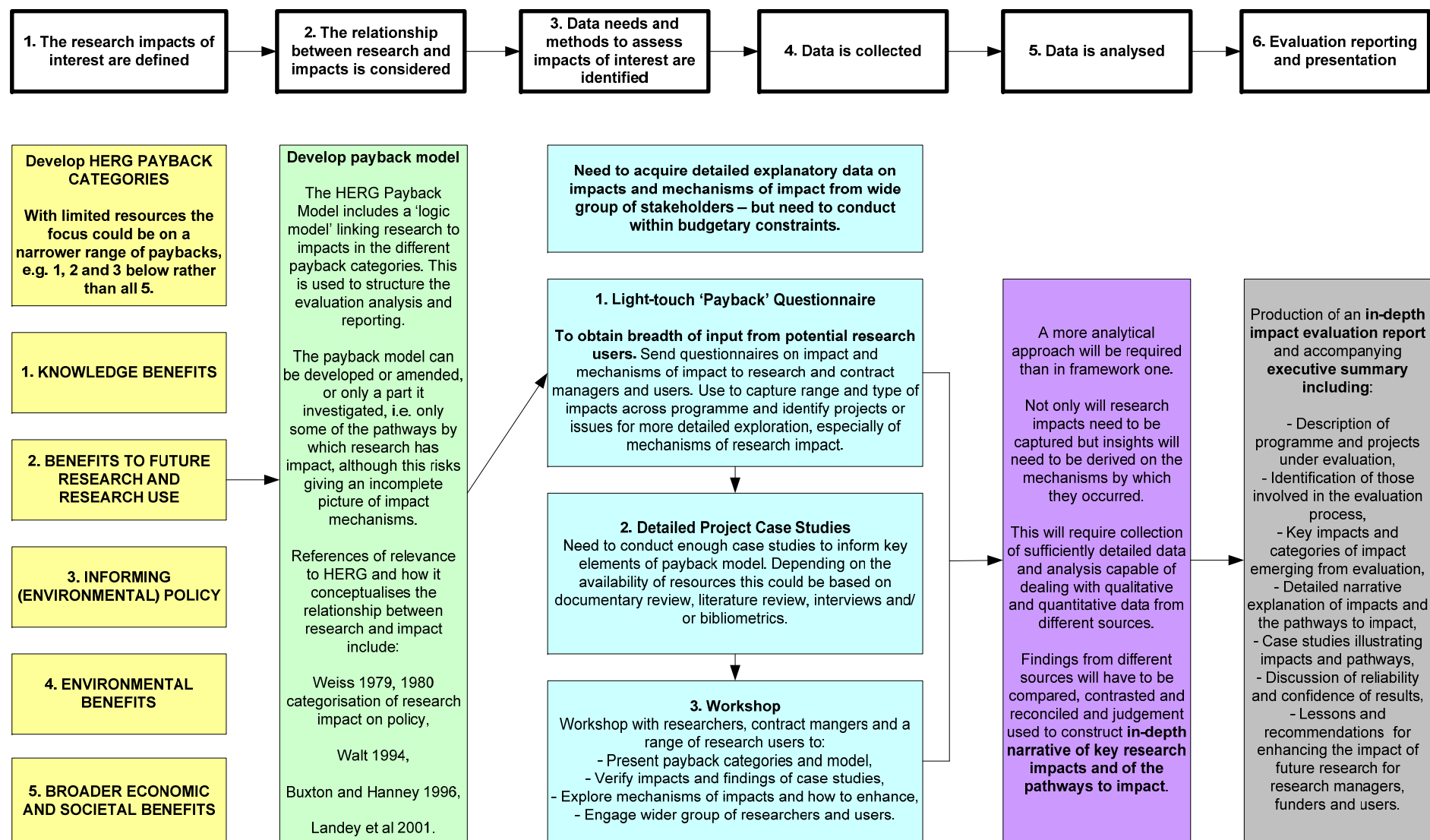
### **Step 7 - Evaluation reporting, dissemination and reporting**

Once the necessary evaluation data has been collected and analysed, the outcomes of the impact evaluation will be reported in a full impact evaluation report and associated executive summary, including headings such as:

- Description of programme and projects under evaluation,
- Identification of those involved in the evaluation process,
- Impacts identified and categories of impacts identified,
- A summary impact score based on ratings of impacts included in the payback questionnaire (although this needs to be built into the evaluation design from outset),
- Detailed narrative explanation of impacts and pathways to impact,
- Case studies illustrating impacts and pathways,
- Lessons and recommendations for enhancing the impact of future research on policy (and other desired impact areas) for researchers, funders and users.

The key findings of this report should be presented and circulated to research teams and research commissioning teams and management. Additional actions will need to be taken to ensure the findings are embedded in the activities of both researchers and funders to ensure that they are acted upon.

**Figure 9. Adaptation of the HERG Payback Model to produce a strategic approach to impact evaluation**



## **Box 2. Summary of a Strategic Impact Evaluation Framework**

### ***When to use this approach:***

This approach can be used when resources are more readily available, for an internally-driven, programme-related impact evaluation carried out by an experienced, independent evaluation lead (perhaps as part of a wider organisational strategic development process). The primary motivation behind the evaluation is to facilitate organisational learning about how to enhance the impact of research, and the effectiveness of existing impact-related activities and institutional structures. The evaluation seeks to explore the ‘enlightenment’ function of research programmes, i.e. the contribution to knowledge rather than direct use, as well as more tangible, instrumental impacts. Up to six months are available for completion of the evaluation, and it is expected to engage a wide array of stakeholders and potential users.

### ***Key evaluation needs:***

‘Essential’ criteria to consider when developing this impact evaluation approach include the ability to: evaluate programmes effectively; provide opportunities for consensus building and organisational learning; engage external stakeholders; explore mechanisms behind research impact, allowing narrative description/explanation; and to produce results with high confidence and rigour. ‘Desirable’ criteria include the ability to: consider short-term to long-term impacts; ensure independence; consider quality of research; and to be conducted at relatively low cost and relatively quickly.

### ***The evaluation approach:***

An adapted version of the HERG Payback Model best meets the needs of this type of evaluation. The adaptations seek to reduce the resource-intensity of the original HERG framework and to help build greater consensus and organisational learning. A revised HERG framework would therefore be conducted by:

- **Refining the impact payback categories** of interest to focus on environment-related impacts rather than health.
- **Completion of light-touch ‘payback’ questionnaire** to obtain breadth of input from potential research users in a relatively low cost manner. This could involve web-based questionnaires to explore insights on impact and mechanisms of impact from researchers, contract managers and research users.
- **Project case studies** to obtain detailed insights into impacts and the mechanisms by which they occur by for a selection of project in the programme of interest.
- **Workshop** with researchers, contract managers and research users to explore and verify pathways and mechanisms of impacts proposed from the analysis and engage a wider group of researchers and users in the process and the use of its outcomes.

### ***Evaluation reporting:***

Upon completion of the data collection and analysis, a full impact evaluation report and executive summary document would be produced which could include: Description of programme and projects under evaluation; identification of those involved in the

evaluation process; impacts identified and described; detailed narrative explanation of impacts and

pathways to impact; case studies illustrating impacts and pathways; lessons and recommendations for enhancing the impact of future research on policy for researchers, funders and users.

The key findings of this report should be presented and circulated to research teams and research commissioning teams and management. Additional actions will need to be taken to ensure the findings are embedded in the research plans and research management activities.

## 8. References and Sources of Further Information

The references cited in this report are given below along with some other references relevant to each the themes below.

### Evaluation generally:

Brutscher, PB., Wooding, S., Grant, J. (2008) 'Health Research Evaluation Frameworks: Technical Report' [online]. Available from: [http://www.rand.org/pubs/technical\\_reports/2008/RAND\\_TR629.pdf](http://www.rand.org/pubs/technical_reports/2008/RAND_TR629.pdf)

COSEPUP (1999) 'Evaluating Federal Research Programs: Research and the Government Performance Results Act. The Summary and Statement of the Problem' [online]. Available from: [http://www.nap.edu/catalog.php?record\\_id=6416](http://www.nap.edu/catalog.php?record_id=6416)

Garrett, JL. And Islam, Y. (1998) 'Policy Research and the Policy Process: Do the twain ever meet?' *IIED Gatekeeper Series 74* [online]. Available from: <http://www.iied.org/pubs/display.php?o=6138IIED>

### Specific issues associated with evaluation:

CAHS (2009) 'Making an Impact: A preferred framework and indicators to measure returns on investment in health research'. Canadian Academy of Health Sciences [online]. Available from: [http://www.cahs-acss.ca/e/pdfs/ROI\\_FullReport.pdf](http://www.cahs-acss.ca/e/pdfs/ROI_FullReport.pdf)

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

Crewe, E. and Young, J. (2002) 'Bridging Research and Policy: Context, evidence and links' [online]. Available from: <http://www.gdnet.org/rapnet/pdf/Bridging%20R&P%20WP.pdf>

Kanninen, S. and Lemola, T. (2006) 'Methods for evaluating the impact of basic research funding: An analysis of recent international evaluation activity' [online]. Available at: [http://www.aka.fi/Tiedostot/Tiedostot/Julkaisut/9\\_06%20Methods%20for%20Evaluating.pdf](http://www.aka.fi/Tiedostot/Tiedostot/Julkaisut/9_06%20Methods%20for%20Evaluating.pdf)

Kivimaa, P, Mela, H, and Furman, E. (2008) 'Approaches and practices in the evaluation of environmental research programmes' Report to the SKEP ERA-NET [online]. Available at: [http://www.skep-era.net/site/files/WP3\\_Evaluation\\_Programmes.pdf](http://www.skep-era.net/site/files/WP3_Evaluation_Programmes.pdf)

Otronen, M. (2004) 'The evaluation of social relevance in an environmental research programme' *Research Evaluation* 13(1): 43-50 (original study full evaluation report available from: <http://www.aka.fi/fibre/final/impact.pdf>)

Raitzer, D. and Ryan, JG. (2008) 'State of the art in impact assessment of policy-oriented international agricultural research' *Evidence and Policy* 4(1): 5-30

## **Particular methods of evaluation:**

### **1. HERG**

CAHS (2009) 'Making an Impact: A preferred framework and indicators to measure returns on investment in health research'. Canadian Academy of Health Sciences [online]. Available from: [http://www.caahs-acss.ca/e/pdfs/ROI\\_FullReport.pdf](http://www.caahs-acss.ca/e/pdfs/ROI_FullReport.pdf)

Hanney, S., Packwood, T. and Buxton, M. (2000) 'Evaluating the benefits from health research and development centres: A categorisation, a model and examples and application' *Evaluation* 2000 (6): 137  
(Available from: <http://evi.sagepub.com/cgi/content/abstract/6/2/137>)

Hanney, SR., Gonzalez-Block, MA., Buxton, MJ. and Kogan, M. (2003) 'The utilisation of health research in policy-making: concepts, examples and methods of assessment' *Health Research Policy and Systems* 1: 2  
(Open access journal available from: <http://www.health-policy-systems.com/content/1/1/2>)

Hanney, SR., Grant, J., Wooding, S. and Buxton, MJ. (2004) 'Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK's 'Arthritis Research Campaign'' *Health Research Policy* 2: 4  
(Open access journal available from: <http://www.health-policy-systems.com/content/2/1/4>)

Kalucy, L., Bowers, EJ., McIntyre, E., Reed, R. (2007) 'Primary Health Care Research Impact Project: Final Report Stage 1' [online]. Available from: [http://www.phcris.org.au/phplib/filedownload.php?file=/elib/lib/downloaded\\_files/publications/pdfs/phcris\\_pub\\_3338.pdf](http://www.phcris.org.au/phplib/filedownload.php?file=/elib/lib/downloaded_files/publications/pdfs/phcris_pub_3338.pdf)

Kwan, P., Johnston, J., Fung, AYK., Chong, DSY., Collins, RA. and Lo, SV. (2007) 'A systematic evaluation of payback of publicly funded health and health services research in Hong Kong' *BMC Health Services Research* 7: 121  
(Open access journal available from: <http://www.biomedcentral.com/1472-6963/7/121/abstract>)

Nason, E., Janta, B., Hastings, G., Hanney, S., O'Driscoll, M., Wooding, S. (2008) 'Health Research – Making an Impact: The economic and social benefits of HRB-funded research' [online]. Available from: [http://www.hrb.ie/uploads/tx\\_hrbpublications/Health\\_Research-Making\\_an\\_Impact.pdf](http://www.hrb.ie/uploads/tx_hrbpublications/Health_Research-Making_an_Impact.pdf)

Ootwijn, W, Hanney, S, Ligtoet, A, Hoorens, S, Wooding, S, Grant, J, Buxton, M, Bouter, L. (2008) 'Assessing the impact of health technology assessment in the Netherlands' *International Journal of Technology Assessment in Health Care* 24: 3

Wooding, S., Hanney, S., Grant, J. (2005) 'Payback arising from research funding: evaluation of the Arthritis Research Campaign' *Rheumatology* 44: 1145-1156

## **2. Research Impact Framework**

Kuruville, S., Mays, N., Pleasant, A. and Walt, G. (2006) 'Describing the impact of health research: a Research Impact Framework' *BMC Health Services Research* 6: 134 (Open access journal available from: <http://www.biomedcentral.com/1472-6963/6/134>).

## **3. Australian RQF**

Armstrong, R., Waters, E., Crockett, B., and Keleher, H. (2007) 'The nature of evidence resources and knowledge translation for health promotion practitioners' *Health Promotion International* 22 (3): 254-260 (<http://www.ncbi.nlm.nih.gov/pubmed/17596543>)

Department for Education, Science and Training (2005) 'Research Quality Framework: Assessing the quality and impact of research in Australia'. Issues Paper [online]. Available from: <http://www.maths.anu.edu.au/other/ncms/rqf-research-quality-issues-paper.pdf>

Department for Education, Science and Training (2006) 'Research Quality Framework: Assessing the quality and impact of research in Australia'. The Recommended RQF [online]. Available from: <http://www.dest.gov.au/NR/ronlyres/7E5FDEBD-3663-4144-8FBE-AE5E6EE47D29/17543/RecommendedRQF2.pdf>

Donovan, C. (2008) 'The Australian Research Quality Framework: A Live Experiment in Capturing the Social, Economic, Environmental and Cultural Returns of Publicly Funded Research' *New Directions for Evaluation* 118: 47-60

## **4. Irish EPA approach**

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

EPA (2009) 'Innovation for a green economy. Environment and Technology: A win-win story' [online]. Available from: [http://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.

## **5. RAPID Outcome Assessment**

Crewe, E. and Young, J. (2002) 'Bridging research and policy: Context, evidence and links' ODI Working Paper 173 [online]. Available from: <http://www.odi.org.uk/resources/download/151.pdf>

Hooten, N., Lee-Smith, D., Nasinyama, G. and Romney, D. (2006) 'Learning lessons from the Kampala urban agriculture policy process' [online]. Available from:

<http://www.pppppc.org/content/files/documents/Kampala%20report%20final%20WEB.pdf>

Hovland, I. (2007) 'Making a difference: M&E of policy research' ODI Working Paper 281 [online]. Available from: <http://www.odi.org.uk/resources/download/1751.pdf>

ODI (*online tool*) 'RAPID Outcome Assessment Handout' [online]. Available from: [http://www.outcomemapping.ca/download.php?file=/resource/files/simonhearn\\_en\\_RA\\_PID Outcome Assessment.pdf](http://www.outcomemapping.ca/download.php?file=/resource/files/simonhearn_en_RA_PID Outcome Assessment.pdf)

## **Evaluation tools:**

### **1. Economic approaches**

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

Mollas-Gallart, J. and Tang, P. (2007) 'Report of the ESRC Impact Evaluation Methods Workshop 20<sup>th</sup> March 2007' [online]. Available from: [http://www.esrc.ac.uk/ESRCInfoCentre/Images/ESRC%20Impact%20Evaluation%20Methods%20Workshop\\_tcm6-24818.pdf](http://www.esrc.ac.uk/ESRCInfoCentre/Images/ESRC%20Impact%20Evaluation%20Methods%20Workshop_tcm6-24818.pdf)

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

### **2. Logic models**

Engel-Cox, JA., Van Houten, B., Phelps, J., Rose, SW. (2008) 'Conceptual Model of Comprehensive Research Metrics for Improved Human Health and Environment'. *Environmental Health Perspectives* 5: 583-592  
(available from: <http://ehp03.niehs.nih.gov/article/action?articleURI=info%3Adoi%2F10.1289%2Fehp.10925/>)

Trochim, WM., Marcus, SE., Masse, LC., Moser, RP., Weld, PC. (2008) 'The evaluation of large research initiatives: a participatory integrative mixed methods approach' *American Journal of Evaluation* 29: 8-28  
(available from: <http://aje.sagepub.com/cgi/content/short/29/1/8>)

### **3. Indicator-based approaches**

CAHS (2009) 'Making an Impact: A preferred framework and indicators to measure returns on investment in health research'. Canadian Academy of Health Sciences [online]. Available from: [http://www.caahs-acss.ca/e/pdfs/ROI\\_FullReport.pdf](http://www.caahs-acss.ca/e/pdfs/ROI_FullReport.pdf)

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

#### **4. Use of a counterfactual**

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

#### **Review papers:**

Boaz, A., Fitzpatrick, S. and Shaw, B. (2008) 'Assessing the impact of research on policy: A review of the literature for a project on bridging research and policy through outcome evaluation' Report to Defra [online]. Available from: [http://www.psi.org.uk/pdf/2008/bridgingproject\\_report\\_with\\_appendices.pdf](http://www.psi.org.uk/pdf/2008/bridgingproject_report_with_appendices.pdf)

Powers, A., Rose, SW., Bernichon, T., Engel-Cox, J., Sill, M., Versendaal, K. (2006) 'Program Assessments and Evaluation for NIEHS: Work Assignment 1 – Literature Review on Research Impact Assessment Methods and Concepts'. Arlington. Centers for Public Health Research and Evaluation.