Assessing the impact of research on policy: a literature review

Annette Boaz, Siobhan Fitzpatrick and Ben Shaw

Understanding the impact of research on policy is a vital, and often overlooked, element of policy-making. A systematic literature review was conducted to examine methods for evaluating the impact of research on policy outcomes. The review focused in particular on strategic policy levels (rather than implementation) and waste, environment and pollution policy. The review draws on an international literature, although it is limited to English language publications. The findings identify the different methods used, the advantages and disadvantages of different approaches and the methods that are most effective (particularly in terms of cost). The field of research impact evaluation is currently experiencing rapid development, and there is scope to develop and apply new conceptual frameworks and innovative methods for evaluation, and to conduct and publish more research impact evaluations.

There is a growing interest in evaluating the impact of research, particularly in understanding how policy-making is evolving and the extent to which policy is directed by varying kinds of research. This interest can be discerned in UK research councils (BBSRC, 2005; Davies et al., 2005), international organisations (Cunningham et al., 2001; World Bank, 2004), the European Union (EC, 2005; Georgiou and Davis, 1988) and academia (Elliott and Popay, 2000, Walter et al., 2004). Furthermore, it can be seen in a wide range of different disciplinary and policy domains. With these developments has come an interest in a wider variety of measures of impact, reaching out to assess the non-academic influence of research investments (Molas-Gallart and Tang, 2007; Research Councils UK, 2007). Recent evaluations have considered both the broad impacts of research and specific impacts on policy.

The factors motivating organisations and individuals to consider research impact are manifold. They include accountability, performance, promotion of organisational achievements, learning and moderating between competing stakeholders/interests (AHRC, 2007; Big Lottery Fund, 2005). In particular, the pressure for accountability is clear in a number of recent documents (BBSRC, 2005; NAO, 2003) and increased government investment in research and development in the UK (HMT, 2002) has been accompanied by a new programme of independent reviews to evaluate the quality and use of science in government departments.

This paper reports on a systematic literature review that aimed to answer the following question: How do you (best) evaluate the impact of research programmes on policy? To answer this question, the following sub-questions were addressed:

- How are the impacts of research programmes evaluated?
- What are the advantages and disadvantages of different approaches?
- Which methods are most effective?
- Which methods offer value for money?

The literature review was conducted as part of a larger research project commissioned by the UK
Department of Environment, Food and Rural Affairs (Defra) (Boaz et al., 2008). The project was intended to review recent developments in the area of evaluation, with particular reference to evaluating the impact of research on policy, and then develop and apply an evaluation methodology to two of Defra’s research programmes: the Waste and Resources Evidence Programme and the Sustainable Consumption and Production Evidence Base.

Structure of the paper

This paper begins with a description of the methods used to conduct the review. The next section describes a mapping exercise that categorises the literature around a number of key variables. While mapping is a relatively descriptive process, it generates insights into the landscape of research impact evaluation. The next two sections outline and discuss what can be learnt from the literature with regard to the review sub-questions, and identify gaps in the literature relevant to these questions. The discussion section returns to the main review question: How do you (best) evaluate the impact of research programmes on policy?

Definitions

There are many different questions an evaluator might ask about a project or programme; ranging from ‘Has it worked?’ to ‘Why and How it has (or has not) worked?’ and whether or not it has been acceptable to participants. The particular focus of this review is the impact of research on policy. Multiple terms are used to describe research impact, including outcomes, benefit, payback, translation, transfer, uptake and utilisation (Beacham et al., 2005; Carden, 2004; Flint, 1998). These terms are often used interchangeably, although there have been some efforts to distinguish between outcomes and impacts. For example, the International Development Research Centre (IDRC) argues that it is helpful to think in terms of behavioural outcomes rather than longer-term policy impacts. These different terms have a shared interest in change that lies beyond the research process and its primary outputs. They are concerned primarily with policy outcomes or impacts rather than outputs or process indicators. Outcomes have been defined as: ‘The end results of a programme for the people it was intended to serve’ (Weiss, 1998: 8).

Although this review searched for and focused on impact evaluations, many of the studies identified also sought to map processes and capture key outputs. These multifaceted, multi-method approaches to evaluation are captured in the review.

The review focused on strategic level policymaking (reflecting the priorities of the UK national government department funding the review) rather than implementation or regulation and as a consequence misses out some impacts of interest. Additional effort was made to identify environmental policy examples, although these were relatively few in number.

Methods

The review used a number of systematic review methods, including identifying a question, developing a well-defined search strategy, using pre-defined select criteria and completing data extraction sheets for each study (Boaz et al., 2001; Petticrew and Roberts, 2005). The main advantage to this approach is that it seeks to be as transparent as possible and therefore open to scrutiny.

Search strategy

The search strategy included three elements: database searches, web searches and expert contacts. The review was international in its scope, but the search strategy only considered sources written in English due to time and funding constraints.

Database searches Ten key databases were searched by an information scientist: INSPEC; Engineering Index; Web of Knowledge (science and technology and social science sections); International Bibliography of the Social Sciences (IBSS); World Political Science Abstracts; British Library Direct (covering the last five years of the INSIDE
From the titles resulting from these searches, 142 papers were identified as relevant, 27 of which were selected for inclusion in the review based on a closer examination of the abstract (or full paper where an abstract was not available). A further search was conducted of three library catalogues: the British Library; SPRU – Science and Technology Policy Research; and the British Library for Development Studies (BLDS). This identified 47 papers based on titles, 29 of which were selected for inclusion in the review based on the abstract.

Web searches The web search involved a more creative search of key organisations, including UK government departments, select committees of the UK Parliament, evaluation societies and research councils. The websites of 30 organisations were searched (see Box 1). The database and web search were conducted in September 2007.

Box 1. The web search

AHRC (Arts and Humanities Research Council), Australian Research Council, BBSRC (Biotechnology and Biological Sciences Research Council), Canadian Health Services Research Foundation, Canadian Environmental Assessment Research Council, CGIAR (Consultative Group on International Agriculture Research), UK Department of Health, Defra (Department for Environment, Food and Rural Affairs), EPSRC (Engineering and Physical Sciences Research Council), ERFF (Environment Research Funders’ Forum), European Environment Agency, ESRC (Economic and Social Research Council), GAO (Government Accountability Office), IDRC (International Development Research Centre), IFPRI (International Food Policy Research Institute) National Academies Press, NSERC (Natural Sciences and Engineering Research Council of Canada), NSF (National Science Foundation), NHS SDO (NHS Service Delivery Organisation), NRC (National Research Council), ODI (Overseas Development Institute), PREST (Centre for Science and Technology Policy and Management Research), REPP (Research Evaluation and Policy Project), RURU (Research Unit for Research Utilisation), SCIE (Social Care Institute for Excellence), SPRU (Science and Technology Policy Research), Technopolis, UKES (UK Evaluation Society), and the World Bank.

Citation tracking and expert contacts Citation tracking involved following up references and contacts emerging from documents identified through the initial stages of the search, and proved to be a particularly fruitful source of relevant material. Manual searches of two key journals in the field (Evidence and Policy and Evaluation Research) also identified some additional papers.

The final part of the search involved writing to six key individuals identified through the review and discussions with the research team, for their feedback and advice on any missing references:

- Fred Carden (International Development Research Centre)
- Luke Georghiou (Centre for Science and Technology Policy and Management Research PREST)
- Steve Hanney (Health Economics Research Group, Brunel University)
- Sandra Nutley (Research Unit for Research Utilisation, University of Edinburgh)
- John Holmes (Department of Earth Sciences, Oxford University)
- John Young (RAPID programme, Overseas Development Institute)

Comments and suggestions for additional papers were received from five of the six. The consultation was particularly useful in identifying new research: Nine of the 10 additional papers were either published in 2007 or were due to be published in 2008.

In total, 351 papers were identified through the different search strategies, of which 156 met the selection criteria for inclusion in the review.

Selection criteria

There is a wide-ranging literature exploring the relationship between research and policy but this review focuses on identifying methods for evaluating the impact of research on policy. The main inclusion criterion was the relevance of the paper or report to the purpose of the review. To this end, papers were included that addressed one of the following:

- An evaluation of the impact of research on policy;

There is a wide-ranging literature exploring the relationship between research and policy but this review focuses on identifying methods for evaluating the impact of research on policy.
Assessing the impact of research on policy

- A reflective piece about evaluation of the impact of research on policy;
- A description of method(s) for evaluating impact of research on policy;
- Evaluation of effectiveness/value for money of methods; and
- A review of methods for evaluating impact more generally.

Data extraction

The details of the literature identified were recorded in Endnote (reference management software), and a data extraction sheet (see Table 1) was completed for each study identified for inclusion. A subset of papers (5%) was read and discussed by all three members of the review team to improve consistency between reviewers and to ensure that the data extraction sheets covered all the critical fields. Following this process, a number of additional fields were added to the data extraction sheet, including details of how the paper was identified, and its disciplinary focus.

Synthesis

The synthesis was organised around the questions identified in the original research specification.

Table 1. The data extraction sheet (compressed)

<table>
<thead>
<tr>
<th>Data extraction sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID reference number</td>
</tr>
<tr>
<td>Method of identification (e.g. database)</td>
</tr>
<tr>
<td>Details of publication</td>
</tr>
<tr>
<td>Author</td>
</tr>
<tr>
<td>Title</td>
</tr>
<tr>
<td>Source (journal/conference, etc.)</td>
</tr>
<tr>
<td>Year/volume/pages/</td>
</tr>
<tr>
<td>Country of origin</td>
</tr>
<tr>
<td>Institutional affiliation</td>
</tr>
<tr>
<td>Field (e.g. environmental research)</td>
</tr>
<tr>
<td>Type of paper (conceptual, empirical, descriptive)</td>
</tr>
<tr>
<td>Research question/aim</td>
</tr>
<tr>
<td>Study design</td>
</tr>
<tr>
<td>Theory</td>
</tr>
<tr>
<td>Analysis</td>
</tr>
<tr>
<td>Findings</td>
</tr>
<tr>
<td>Paper includes:</td>
</tr>
<tr>
<td>- An evaluation of the impact of research on policy;</td>
</tr>
<tr>
<td>- A reflective piece about evaluation of the impact of research on policy;</td>
</tr>
<tr>
<td>- A description of method(s) for evaluating impact of research on policy;</td>
</tr>
<tr>
<td>- Evaluation of effectiveness/value for money of methods;</td>
</tr>
<tr>
<td>- A review of methods for evaluating impact more generally</td>
</tr>
<tr>
<td>- Others of relevance (give reason)</td>
</tr>
</tbody>
</table>

Mapping the literature

The first stage of the analysis involved mapping and categorising the literature around a number of key variables:

- Topic;
- Country of origin;
- The way in which the papers and reports addressed the review question; and
- The methodologies used or advocated for conducting impact evaluations of research on policy.

The majority of the papers were identified through the internet search and from the reference lists in papers and reports. Personal recommendations also made a strong contribution. Although only 27 papers were identified through the database searches, this approach did provide access to papers not identified through other sources, reinforcing the value of a mixed-method search strategy.

This review particularly focused on evaluations of the impact of research on policy. Of the 156 studies, 58 reported on empirical studies of the relationship between research and policy, of which 42 were specific evaluations of the impact of research on policy. Most of the remainder were reflections on, or descriptions of, methods or approaches to impact evaluation. The review identified a wide-ranging literature, reflecting a cross-sectoral interest in the issue of research impact on policy. In particular, there are large health and international development literatures as well as a generic literature on utilisation and evaluation.

A wide range of research methods and approaches have been applied to the study of research impact (Figure 1). The most frequently mentioned and used are semi-structured interviews, case studies and documentary analysis. This is consistent with another recent review of impact studies for the UK Health Technology Assessment programme (Hanney et al, 2007). However, most papers mention more than one method. In terms of the use of methods, in empirical studies, it is interesting to note that while bibliometric analysis was frequently mentioned as a method, it was rarely used to measure the impact of research on policy. By contrast, telephone interviewing was used almost as frequently as it was discussed. There are many more references to, and uses of, qualitative methods than quantitative methods.

Many different frameworks for interpretation are discussed in the literature but only a small number are actually used in impact evaluations (Figure 2). In
particular, RAPID Outcome Assessment, the HERG Payback Model and economic analysis are applied in empirical studies (although the last normally forms part of a wider analysis). The frameworks are usually applied exclusively by the individuals and organizations that developed them.

Findings

Although the review focused on identifying approaches to assessing the impact of research on strategic, national-level policy, many of the studies included also sought to capture the complexity of processes leading to impact. Where these approaches were used as part of impact evaluations they are also captured in the review.

*How are the impacts of research programmes evaluated?*

The literature (Merkx *et al.*, 2007; Ryan, 2002) argues that evaluators should be driven by the need to identify methods which are the most ‘fit for purpose’. This is not a straightforward task. Some methods are very flexible and lend themselves to a wide variety of evaluative tasks. For example, semi-structured interviews can be used to track impact...
Assessing the impact of research on policy

A considerable amount of impact evaluation is conducted in the field of international development. The emphasis seems to be on qualitative, participatory evaluations with a focus on learning and service improvement. Most of the evaluations use case studies or mixed-method approaches (Acosta and Douthwaite, 2005; Christiansen with Hovland, 2003; Flint, 1998; Leksmono et al., 2006) and participation is encouraged, primarily through workshops. While engaging potential users in the process of impact evaluation is both complex and time-consuming, Carden argues that it has the potential to increase both relevance and use of the study (Carden, forthcoming).

A further distinction between evaluation in international development and other domains is the use of field visits, where evaluators observe practice on the ground (Flint, 1998; Thiele et al., 2006), and ‘storytelling’ to explore and explain impacts (Acosta and Douthwaite, 2005; Sayce and Norrish, 2006). Participatory research can generate valuable insights, particularly when triangulated with data from other sources. However, the reliability of the more qualitative data collected in both international development evaluations and in other fields seems to vary. While stories can be based on research evidence, they can also be largely anecdotal and highly subjective. The focus on constructing positive utilisation narratives (Acosta and Douthwaite, 2005; Davies and Dart, 2005; Hovland, 2007) strays dangerously close to the line between evaluation and promotion.

Whereas impact evaluations in international development pose important questions — about who conducts the evaluation (professional evaluators, programme staff, service users), the involvement of stakeholders in the process and the value of a mixed-method, qualitative approach — conceptual literature in the United States reflects the requirement outlined in the US Government Performance and Results Act for a greater use of quantitative methods and performance indicators (Arnold et al., 2005).

In particular, commentators have struggled with the applicability of modelling and economic analyses to the evaluation of research impact (Smith, 1998). For example, Zilberman and Heiman (1999) developed an economic formula to assess the net benefit of research that led to innovations, but acknowledged the empirical challenges of obtaining actual estimates. They suggest that the most important contribution of impact assessments of research on policy may be preventative (i.e. to screen wasteful policy proposals). The lack of evidence available (Zilberman and Heiman, 1999), the quality of existing data and the uncertainties characteristic of the research process (Kostoff, 1995) make methods such as cost–benefit analysis (used so widely in industry to look at the costs involved and the benefits generated) extremely difficult to apply (Hanney et al., 1999). Empirical evaluations conducted in the USA tend to continue to rely on a mixed-method approach to overcome the limitations of individual methodologies by a process of triangulation (Kostoff, 1994).

The main method used for evaluating the impact of European Union (EU) research programmes is panel reviews (Arnold et al., 2005). There have been a number of recent reviews exploring alternative methods for evaluating EU investments such as bibliometrics, econometrics and social analysis (Fahrenkrog et al., 2002). However, the value of panels continues to be emphasised and the suggestion is that other methods would be used as a...
Although panel reviews attract criticism for their reliance on experts, they do have the advantage of building ownership through the participation of key individuals from different countries within the European Union complement rather than a replacement. Although panel reviews attract criticism for their reliance on experts, they do have the advantage of building ownership through the participation of key individuals from different countries within the EU. This is particularly important in the European context and underlines the importance of selecting methods that are both fit for purpose and appropriate to the needs of key stakeholders. While panel reviews are the predominant approach in the EU, other methods have been applied. One example is a large-scale internet survey of researchers funded as part of the EU Framework Programme 5 and conducted in 2004 (Guy et al, 2005).

A European Commission report on different approaches to evaluating energy research in EU countries (EC, 2005) identified a shift in practice away from project management assessments (e.g. project delivered on time and to budget) to an interest in the impacts on science, policy and industry. While the paper described diverse practices within Europe (e.g. in Denmark there is a focus on evaluating the use of research within target groups), the study also identified common interests across the EU, including the need to demonstrate the additionality of research investments and a commitment to developing meaningful quantitative indicators.

Practice varies considerably between different countries and policy contexts. There are common themes in terms of the interest in developing and applying new methods, such as network analysis (Almeida and Bascálo, 2006; Fahrenkrog et al., 2002; Kostoff, 1994; Lewis, 2005); new understandings of changes such as the increased focus on evaluating behaviours (Earl and Carden, 2002); the need to develop existing methods to ensure they are fit for purpose, such as new developments in citation analysis (CHASS, 2005; Nederhof, 2006); and the promotion of mixed-method approaches. However, for evaluation to be relevant and usable in different contexts some diversity is likely to remain in evaluation practice.

Table 2. Approaches and methods for gathering data

<table>
<thead>
<tr>
<th>Approach/method*</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Semi-structured interviews | A flexible interview around a framework of themes, with pre-identified key actors | - Flexible structure enables interviewees to respond in own terms, and interviewer to respond as part of a two-way conversation  
- Framework ensures comparability of interviews  
- Seen as most appropriate when unravelling diverse layers and subtle nuances | - Issues of bias and attribution  
- Training for interviewer is necessary (to avoid pre-empting answers or not probing enough)  
- Time-intensive (collection and analysis)  
- Interviewees may themselves be unaware of indirect influences of research  
- Time recall  
- Difficulty in selecting cases  
- Issues of bias and attribution  
- Often over-emphasises importance of research – ‘supply perspective’  
- Time-intensive to ensure rigour  
- Difficult to apply a common framework across case studies  
- Relies upon the quality of existing records and access to these  
- No single methodology for analysis |
| Total = 57 | | | |
| Case study analysis | An empirical approach that explores in depth a project/programme, describing and explaining how and why developments of interest have occurred | - Can be descriptive and explanatory  
- Can demonstrate pathways from research to impact  
- Potential to combine sources and methods (triangulation)  
- Explores context  
- Can be applied to a wide range of sources (including policy statements, technical reports, minutes, speeches) | - Time recall  
- Difficulty in selecting cases  
- Issues of bias and attribution  
- Often over-emphasises importance of research – ‘supply perspective’  
- Time-intensive to ensure rigour  
- Difficult to apply a common framework across case studies  
- Relies upon the quality of existing records and access to these  
- No single methodology for analysis |
| Total = 56 | | | |
| Documentary analysis | Exploration and interpretation of existing documents. Can elicit quantitative or qualitative findings. Often used in conjunction with other methods | - Provides contextual understanding  
- Potentially cost-effective | | |
| Total = 45 | | | |

*Includes in-depth reviews, expert interviews, focus groups and other methods.

What are the advantages and disadvantages of different approaches?

The wide range of approaches used to collect data on the impact of research on policy is listed in Table 2. Their advantages and disadvantages are also included in the table, along with a short description of each method. These can be grouped as follows:

- Qualitative methods including semi-structured interviews, documentary analysis, field visits and observations. These methods are used to generate rich descriptive and explanatory data that can be used to look at both utilisation pathways and the policy context in which utilisation is or is not.
### Table 2. (continued)

<table>
<thead>
<tr>
<th>Approach/method*</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Bibliometrics/citation analysis**  
Total = 37 | Method for quantifying the impact of research by counting the number of outputs and citations, and analysing citation data | • Suitable for repeated analyses and comparisons  
• Measures original research, not the programmes resulting from it  
• Identifies research deemed to be important by subsequent research  
• Traditionally considers academic citations, but extended analysis includes grey literature | • Measures outputs not outcomes  
• Not comparable between disciplines  
• Quantity of output may not reflect quality; risks distorting importance of select publications  
• Limited role where publication is not the goal of research  
• Many uncertainties, and thus can only be used as a partial indicator  
• Time lag between publication and citation  
• Critical as eliciting acceptability rather than validity of findings  
• Qualitative findings only  
• Requires comprehensive information for reviewers | |
| **Peer/panel review**  
Total = 32 | 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 | • ‘Experts’ confer status, credibility, and acceptability on findings  
• Can offer range of constructive feedback to guide process  
• Can be conducted at any time during the process of impact analysis  
• Relatively cost-effective  
• Flexible | • Reflects upon rather than measures impact  
• Time-consuming method, particularly for the experts involved  
• Issues of objectivity and variability (including the quality of experts)  
• Impractical to evaluate a broad area (number of peers involved)  
• Unresponsive to unforeseen issues  
• May require follow-up interviews to fully understand the results | |
| **Surveys**  
Total = 30 | A pre-formatted series of questions asked of multiple actors, generating both quantitative and qualitative data | • Can identify outputs/outcomes associated with particular research  
• Inexpensive means of providing an overview from a range of actors  
• Wider range of stakeholders than possible with interviewing  
• Can identify aspects to focus on in interviews | • Relies upon access to respondents  
• Reflects the bias of those surveyed and those who respond  
• Unresponsive to unforeseen issues  
• May require follow-up interviews to fully understand the results | |
| **Workshop, focus group**  
Total = 20 | An organised discussion with a group of individuals. The groups can involve a range of different stakeholders | • Can be conducted at any time during the process of impact analysis  
• Less expensive than surveys  
• Exploratory in-depth insights  
• Can reach a consensus that individual responses may not | • Risk of sample bias  
• Selective memory of participants  
• Unsuitable for competitive or sensitive topics | |
| **Literature review**  
Total = 18 | Synthesis of existing research relevant to the study. In the context of impact evaluation, usually used with other methods | • Useful initial research to define the scope of an impact study/look at evaluation methods  
• Cost-effective  
• Can ensure information is up-to-date  
• Direct observation of activities | Depends upon ability to identify and access existing research  
• Time-intensive and costly (requires planning and ex-post evaluation if it is to be beneficial) | |
| **Field visit**  
Total = 13 | 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 | | |
| **User evaluations**  
Total = 8 | Participatory method for assessing stakeholder (either users of research or producers of research) satisfaction. May involve interviews, email or telephone survey | • Looks at both research and research utilisation from a stakeholder perspective  
• In-depth understanding of utilisation processes | • Risk that stakeholders have a vested interest in expressing satisfaction  
• Can be both time-intensive and costly | |
| **Telephone interviews**  
Total = 8 | Usually semi-structured interviews, often used as preliminary means of identifying key stakeholders | • Can elicit open-ended information early on in research process  
• Enables greater sampling dispersion  
• Cost- and time-effective  
• Policy-oriented approach to tracking events and processes  
• Can be used to explain ‘how’, ‘what’ and ‘why’ | • Less able to develop rapport with unknown interviewee (restricts line of questioning)  
• Interview length is limited  
• Cannot use visual prompts/sources  
• Relies upon the quality of, and access to, existing documentation  
• Difficult to attribute causality  
• Does not account for indirect impacts  
• Difficult to compare between disciplines  
• Time lag after publication  
• Most commonly used to evaluate the wider impact of research on industry | |
| **Historical tracing**  
Total = 5 | Tracing backwards from an outcome to identify contributing factors, using a range of (usually qualitative) data collection tools | • Can be used to explain ‘how’, ‘what’ and ‘why’ | |
| **Patents/new technologies**  
Total = 6 | Where research may have patentable outcomes, this approach gathers data about the number and nature of patents | Useful to identify linkages between research and specific outcomes | |
Table 2. (continued)

<table>
<thead>
<tr>
<th>Approach/method*</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network mapping and analysis</td>
<td>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.</td>
<td>Reflects complex interactions of the realities of policy-making</td>
<td>Can identify linkages but cannot identify impacts/outcomes as a consequence</td>
</tr>
<tr>
<td>Positive utilisation narratives</td>
<td>A participatory method to identify stakeholder accounts of impact, and the use of secondary analysis to understand who contributed to change and how</td>
<td>Identifies unexpected change</td>
<td>Anecdotal accounts and within-organisation evaluation prone to bias</td>
</tr>
<tr>
<td>Impact log</td>
<td>A means of logging real-time direct impacts and uptake of research (i.e. from informal feedback, field observations)</td>
<td>Quick and cost-effective method of evaluation</td>
<td>Does not identify negative impacts or the non-use of research</td>
</tr>
<tr>
<td>Tracing post-research activity</td>
<td>Follows the impact of research according to channels of diffusion through networks and post-research activity of researchers</td>
<td>Self-evaluation and learning</td>
<td>Cannot be conducted retrospectively</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Records impacts over time</td>
<td>Much of the information is anecdotal and subjective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can be used to construct actor matrices</td>
<td>Does not measure indirect impact</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difficulty in tracking activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Anecdotal</td>
</tr>
</tbody>
</table>

Note: * The totals refer to number of studies making reference to the method.

taking place. However, they are often time- and resource-intensive, and it can be difficult to generalise their findings.

- **Quantitative methods including surveys, bibliometrics and patent/new technology tracking.** These methods are used to quantify the outputs of research. They are suitable for repeat analyses and comparisons, can be used to manage large amounts of data and can be very cost-effective. However, they are often difficult to use in the analysis of research impact on policy. Survey response rates can be poor and bibliometric data usually focus on the quantity (rather than quality) of research outputs rather than outcomes.

- **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 to be used to enhance the quality of the research rather than to assess policy impact. Furthermore, panels are open to accusations of lack of objectivity and are heavily reliant on the quality of their membership.

- **Workshops and focus groups.** These are interactive, consensus-building approaches, but are not suitable for all topics (e.g. sensitive topics that generate polarised views), and they can lack rigour and objectivity.

- **Process tracking including historical tracing, positive utilisation narratives, tracing post-research activity and impact logs.** These methods focus on exploring the utilisation pathway, including how and where research has been used over time. They rely heavily on the quality and relevance of existing records (or record keeping in the case of impact logs), documents and recall of participants in the process.

- **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.

- **Network mapping and analysis.** This approach seeks to construct realistic accounts of the nature of relationships between knowledge and policy. Other research methods are used to generate the data required for constructing maps of the networks that exist between individuals and organisations.

The individual methods listed above are rarely used in isolation. For example, case studies usually involve a combination of qualitative and sometimes quantitative methods to explore utilisation (Michelson, 2006; Norton and Alwang, 1998). Most of the papers identified in this literature review promote the idea of mixed-method approaches to exploring research impact, whether or not as part of a case study (Anderson, 2006; Zilberman and Heiman, 1999). In particular, they argue for a combination of qualitative and quantitative approaches, acknowledging the different contributions the two broad approaches can make.

In addition to the selection of methods, there is some debate about who should be involved in conducting the evaluation. For example, user evaluations involve stakeholders in producing the study, while expert and peer panels also represent a more interactive, consensus-building approach. These approaches have the advantage of building ownership and fostering individual and organisational learning. However, while these approaches can be
Assessing the impact of research on policy

cost-effective, there can be challenges in terms of quality, objectivity and bias, although it is important to note that not all external reviews are objective just as not all internal reviews are subjective. In a review of methods for assessing research impact, Hovland (2007) suggests that a strong evaluation would include elements of both internal and external review. The potential benefits of a mixed approach are echoed elsewhere for practical and methodological reasons. For example, it has been argued (Wooding et al., 2007) that internal evaluators would be in a better position than external evaluators to keep a record of impacts both in the short term and over a longer time period. External evaluators could then focus on a critical point in time when impacts might be expected. Using both internal and external evaluators can also bring different perspectives to bear on the evaluation (Carden, forthcoming).

Most of the identified evaluation approaches can be used for either forward tracking from a piece of research to the outcome or backward tracking from an outcome such as a policy change or document back to the research, although the former (NHS SDO, 2007) is more common than the latter (Hilderbrand and Simon, 2000). Hanney et al (2007) argue that tracking forwards tends to identify a greater level of impact than tracking backwards. They suggest that this might be due in part to the reliance on self-reported data from lead investigators in forward-tracking studies. However, there are some examples where evaluators have sought to track in both directions to construct a rich account of the relationship between research and policy. For example, the RAPID framework developed and applied by the Overseas Development Institute (Hooton et al., 2006; Kostoff, 1995; Leksmono et al., 2006) draws upon episode studies of specific policy changes, case-study analysis of specific research projects, and outcome-mapping approaches. Information is collected using a range of methods including: literature review and documentary analysis; workshops with key staff, partners and actors; stakeholder interviews to triangulate workshop outputs; field visits; and a debriefing session to discuss initial findings.

One of the challenges of applying standard research methods is that they have not been developed for the particular purpose of evaluating the impact of research on policy. The same can be said for some of the frameworks for interpreting data, such as the various approaches to economic analysis, outcome mapping, benchmarking, simulations, episode studies and social analysis. A number of project management tools do include evaluation as a critical stage in the life cycle of a project. These include log frames (Fisher and Holland, 2003; Pielke, 1995), ROAMEF (rationale, objectives, appraisal in detail, monitoring, evaluation and feedback) (Cunningham et al., 2001; Georgiou and Davis, 1988) and outcome mapping (Garfinkel et al., 2006; Leksmono et al., 2006). The key strength of these approaches is that evaluation is embedded within the project management process, thus ensuring that it is not seen as a separate, stand-alone activity, and is considered right from the start of project planning. However, these different approaches do not provide detailed guidance on how to capture learning about processes and how to measure complex outcomes. Thus, despite their useful elements, there are considerable challenges in applying these tools to the evaluation of research impact on policy.

Adapting and/or adapting a conceptual framework or model of the relationship between research and policy is, however, considered to be a valuable step in the development of a plan for studying the impact of research upon policy (Molas-Gallart and Tang, 2007), and a number of models have been developed specifically for this purpose. These include the Health Economics Research Group (HERG) payback model, RAPID outcome assessment and the Research Impact Framework (RIF). For example, the HERG payback model explores five different impact domains:

- Knowledge production;
- Research targeting and capacity building;
- Informing policy and product development;
- Health (and health sector) benefits; and
- Broader economic benefits.

There is much to be learnt from this approach, particularly as it has now been applied to a wide range of policy fields, such as health services and labour market research (Hanney et al., 2004; Wooding et al., 2007). The application of revised citation analysis techniques and a mixed-method approach, combined with a clear conceptual framework (often used alongside a logic model to assess outputs and outcomes), makes the approach attractive as a tool of impact evaluation in different countries and contexts. The RIF assesses policy impact around five dimensions:

- The level of policy-making;
- The type of policy;
- The nature of policy impact;
- Policy networks; and
- Political capital.

Adapting a conceptual framework is considered to be a valuable step in the development of a plan for studying the impact of research upon policy, and a number of models have been developed specifically for this purpose.
This framework generates brief, one-page impact narratives that could be more attractive and applicable in applied research environments (Kuruvilla et al, 2007).

Which methods are most effective?

Only 14 studies in the review reflect directly on the effectiveness of different methods for evaluating the impact of research, and in only one of these was the effectiveness of different approaches a central concern. Hanney et al (1999) compared survey methods and case studies as methods for evaluating the impact of research investments. They conducted an empirical comparison of questionnaires and case studies, and found that questionnaires could provide a relatively robust approach to monitoring a broad research portfolio, but would benefit from a small number of accompanying case studies that would help to verify the information given and to test and improve the questionnaire for future use. This analysis was repeated with similar results as part of an evaluation of the impact of the NHS Health Technology Assessment (HTA) programme (Hanney et al, 2007). The comparative study (Hanney et al, 1999) emphasises the need to develop realistic approaches that fit with the scope of the evaluative task and the resources available.

In the remaining 12 papers, observations were generally limited to short paragraphs of text in background or conclusions sections. These papers are consistent in their analysis, concluding that to gain meaningful insights into research impact, evaluators need to adopt in-depth, exploratory methods such as semi-structured interviews and case studies (Wooding et al, 2004). Williams and Rank (1998) consider case studies to be the best method for evaluating the impact of research, while Jones et al (1999) conclude that there is no substitute for interviewing individuals, particularly within the policy arena. However, while agreeing that evaluators need to go beyond the use of surveys and quantitative indicators in appraising research impact, Anderson (2006) argues that in-depth case studies can be prohibitively expensive.

Kilpatrick et al (1998) consider the costs and benefits of a range of approaches to measuring research impact, noting that methods have been appropriated from evaluations in other fields or from more general social science research. They find that none of the ‘borrowed’ methods completely meets the theoretical needs of analysts assessing research impact, and conclude that while use of a combination of methods may overcome individual shortcomings, this is likely to be a more costly option. Other commentators also warn against the use of single methods, in particular, bibliometrics and econometrics (CHASS, 2005) and argue that there is a need to find new (Molas-Galart and Tang, 2007) or ‘hybrid’ (Wooding et al, 2004) methodologies.

Which methods offer value for money?

The review found very little data on the value for money offered by the different approaches identified to assess the impact of research on policy. At the most, papers include short paragraphs reflecting on the methods used and their acceptability to participants. Typically (as with other research), the costs of conducting impact studies are not reported. The general theme emerging from reflective accounts seems to be that the cheaper and most practicable approaches, such as surveys, are the least likely to yield insights into the process of research utilisation (particularly when conducted in isolation). There is a suggestion in one study (Cranfield and Ward, 2002) that telephone surveys might offer a cost-effective ‘middle ground’ between postal surveys and in-depth interviews: participants concluded that telephone interviews had offered an efficient use of their time as respondents, and had additional value as a method for identifying individuals for in-depth interview.

Two papers (Davies et al, 2005; Hanney et al, 1999), in particular, conclude that it is unlikely that there will be time and resources to consider the impact of a whole portfolio of research. Evaluators will have to consider both issues of sampling (i.e. which research to focus on from within a programme) and the costs of potentially appropriate methods. For example, Hanney et al (1999: 189) conclude that while in-depth case studies generate valuable insights, they are ‘inevitably fairly resource intensive and would not be appropriate for extensive surveying or regular monitoring of an R&D portfolio’.

Beacham et al (2005) also suggest that the comprehensive approach supported by many commentators may well be too complex and expensive relative to the potential benefits, given the intricacy of the evaluative task. In this respect, it is perhaps worth keeping in mind some of the quicker, well-established and low-cost methods, such as using an expert panel to provide an informed opinion about impact, developing a file to record personal testimonies, or developing proxy measures of impact. Davies and Dart (2005) suggest that the ‘most significant change’ technique, which focuses on collecting stories at the field level for discussion by a panel and collation in a report, is worth consideration.

Overall, the frequent recommendation that a variety of evaluation methods should be used is likely to add to the cost of impact studies. The suggestion that new methods need to be devised also has cost implications, given that methodological development work is time- and resource-intensive.

What the literature doesn’t tell us

The review found very few papers focusing on the costs and effectiveness of different approaches to evaluating the impact of research on policy. Only

Science and Public Policy May 2009

265
one study focused on comparing the effectiveness of two popular methods (case studies and survey methods). However, many authors were more reflective about the advantages and disadvantages of different approaches.

The literature also identified fewer than expected empirical evaluations of the impact of research on policy. It is unclear whether this reflects a low level of activity, a failure to publish evaluations, or the reporting of much of this activity in the grey (non-academic, report) literature that can be difficult to identify. The extensive internet search conducted as part of the review did provide access to grey literature, so it is possible that low numbers and lack of publication of evaluations are more likely explanations. Those concrete examples of evaluations that were identified generated powerful insights relevant to this review and it is anticipated that this literature is likely to grow in the future. At present, the majority of studies reflect on the relationship between research and policy rather than the impact of research on policy, and although this is an interesting and relevant literature, it was not the focus of this review.

The relationship between knowledge and decision-making has been a matter of interest for hundreds, if not thousands, of years (Shadish et al, 1991). Much can be learnt from the wider literature exploring the relationships between knowledge and policy in more general terms, and this has been summarised in a series of reports and papers published by the Research Unit for Research Utilisation (Davies et al, 2005; OSI, 2006; Walter et al, 2003, 2004).

Finally, while there is considerable reflection in the literature on the need to develop and test new methods for evaluating the impact of research on policy, or to modify methods used in other contexts, there are fewer examples of attempts to develop and apply novel methodologies. To some extent, new methods are likely to be needed to reflect changing conceptual models of the relationship between knowledge and policy. For example, Gabbay et al (2003) and others have mapped the relationship between knowledge and practice as a mass of circles and boxes connected by a complex web of lines and arrows, which reflects the more realistic view of the non-linear processes of research influence and uptake and represents a world apart from the neat circular and linear ROAMEF and log frames models. Innovations in technology such as mapping software make it possible at least to attempt to capture the complexity of policy-making on paper. This is a field in transition and there may well be applications of innovative methodological practice in development.

**Discussion**

**What is meant by impact?**

The review identified multiple terms to describe research impact, including outcomes, benefit, payback, translation, transfer, uptake and utilisation. However, the studies identified shared a commitment to identifying meaningful approaches to capturing impact and an appreciation of the complexity and diversity of research use.

With an interest in impact come problems with attribution (see Figure 3). In this context, the key question for evaluators is the extent to which impacts are attributable to specific research outputs. It is important to note that the process by which the research is done might have an impact on policy. For example, collaborative research projects may lead to
research having an impact prior to the production of research outputs.

Many approaches remain primarily concerned with the quantity and quality of research outputs rather than outcomes (Arnold et al., 2005; Butler and McAllister, 2007; Buxton et al., 2000). For example, bibliometrics are used to quantify research publications, while peer-review panels consider the quality of the research produced by projects and programmes, often in relation to generic indicators of research quality rather than in terms of its fit to wider programme objectives (Butler and McAllister, 2007; Luukkonen-Gronow, 1987). Attribution becomes more difficult as the evaluator moves away from the initial outputs, to consider the various different types of impact within the policy process. This has been described as going ‘through the door’ rather than ‘to the door’ of the policy-maker (Wooding described as going ‘through the door’ rather than ‘to the door’ of the policy-maker (Wooding et al., 2007). Only a few studies have focused on how policy is made ‘behind the door’ and how research might contribute to that process. For example, one study distinguishes between the use of research conclusions and research recommendations (Brofoss, 1998).

Some studies have promoted a shift in language, preferring to focus on influence rather than impact. Carden (2005) argues that by the time impact has occurred the research is intertwined with too many other factors for it to be possible to ascertain attribution. He encourages evaluators to focus on changes in behaviour as indicators of broader change. Nutley et al makes a similar point:

Research use may thus be more about transformation than straightforward application. There may, however, come a point where such refinements are so extensive that it is no longer legitimate to refer to this process as ‘research use’ at all. (Nutley et al., 2007: 59)

It is perhaps not surprising that the task of quantifying and tracking research outputs (using techniques such as basic bibliometrics) has often been considered to be a more deliverable, and thus attractive task, than assessing the impact of research outputs. Recent developments in the practice of documenting and assessing citations have, however, improved the potential contribution of bibliometrics. For example, one study has attempted to track citations in policy and professional guidance and the mass media (NHS SDO, 2007), and there is a call for more sophisticated analysis to move beyond, for example, approaches in which all impacts are awarded the same value (Woolding et al., 2007).

The evaluation process should be driven by what is meaningful in terms of outcomes, not by what is measurable. Although this review focused on the difficult task of measurement of outcomes, methods for evaluating processes were also included, as it is clear that understanding outcomes requires an exploration of processes, barriers and facilitators as well as the end results.

Recent developments in the practice of documenting and assessing citations have improved the potential contribution of bibliometrics; there is a call for more sophisticated analysis to move beyond, for example, approaches in which all impacts are awarded the same value

How do you (best) evaluate the impact of research on policy?

Although the literature is thin on empirical examples, it is full of reflective pieces. This is a field of practice currently experiencing rapid development, and examples of innovative practice have recently been published (Bumgarner et al., 2006; Delanghe and Muldur, 2007; Wooding et al., 2007). For example, technological developments will now support more complex approaches to bibliometrics and network analysis. Email (Kivistaa et al., 2007) and the web (Guy et al., 2005) offer alternative vehicles for contacting and surveying participants.

Eight key dimensions of best practice emerge from the analysis of the literature undertaken for this study. These are phrased as eight questions that might be addressed by organisations and individuals considering or planning an evaluation to identify the most effective approach to evaluating the impact of research on policy (see Box 2). In particular, they were designed to address the very practical concerns of the funding organization facing a new challenge in terms of evaluating the impact of its research and development programmes. The literature stresses that the earlier questions related to research impact evaluation can be addressed in the research process, the better the evaluation will be (Merkx et al., 2007; Rider Smith with Sutherland, 2002). For example, good quality systems need to be in place to capture and track the use of research.

Practical constraints in terms of time, budget and skills will inevitably guide any evaluation of research impacts on policy. The usefulness of the eight dimensions to those planning and conducting research impact evaluations are the subject of an ongoing study with environmental research funders in Europe.

Conclusion

The demand to demonstrate in a more rigorous way that investments in research have impact on policy is growing. A National Audit Office report found:
Box 2. Eight key dimensions of best practice

1. What conceptual understanding of the relationship between knowledge and policy underpins the evaluation?
2. What are the outcomes or impacts of interest?
3. What evaluation methods might be used to explore the outcomes of interest?
4. How does the evaluation address issues of the attribution of impacts to specific research activities?
5. What is the direction of travel for the evaluation? Backwards from policies and policy-makers; or forwards from research reports, research programmes and researchers; or some combination of the two?
6. Will there be scope to involve stakeholders in the evaluation?
7. Will the evaluation methods selected capture the context and complexity of the research utilisation pathways, helping to understand how, as well as whether, change has occurred? Will they capture changes in behaviours and relationships?
8. Does the timing of the evaluation offer enough time for research impacts to have occurred, but not so much time that recall will be difficult for respondents? Is the evaluation conducted after a policy change has occurred (ex post) or is it predictive of potential use (ex ante)?

Evaluation of the quality of the research process is well established. However, there is a strong and developing emphasis on evaluation to encompass research relevance and value for money, as the link between research results and policy formulation increasingly becomes the focus of attention. (NAO, 2003: 2)

The evaluations identified in this review are consistent in their use of mixed-method approaches, often using a combination of different qualitative methods such as in-depth interviews and documentary analysis. Approximately half of these were explicitly informed by a conceptual framework or model of the research-policy process. The descriptive and reflective papers argue that evaluation in this field could be done better. They advocate the development and application of conceptual frameworks (including, for example, HERG and RAPID but also encouraging the development of new models) and innovative methods for evaluation (such as advanced bibliometrics, developments in survey methods, network mapping and again encouraging the development of new approaches).

Future evaluations might also focus on assessing the impact of interventions designed to promote research use in policy-making, such as knowledge brokers, networks, and broader linkage and exchange programmes. The question: Does current investment in research make a difference? is a challenging one. However, the indications from the literature are that progress in this area is now accelerating and this has the potential not only to improve understanding of the impact that research has on policy, but consequently also to give insights that could make for a more effective interaction between policy and research.

Acknowledgements

This research was funded by the Waste and Resources Evidence Programme of the UK Department for the Environment, Food and Rural Affairs (Defra) as part of a project entitled Bridging Research and Policy through Outcome Evaluation. A full report of the literature review including detailed appendices on the methodology and data extraction sheets for all the studies included in the review can be accessed via: <http://www.psi.org.uk/pdf/2008/bridgingproject_summary.pdf>

The authors would like to thank the experts who contributed to the literature reviewing process (in particular Sandra Nutley who commented on a draft of this paper) and the commissioning team at Defra who provided helpful comments throughout. Finally, thanks to Alan Gomersall for conducting the database searches and to Lesley Grayson for additional support with searching and editing.

References


Butler, L and I McAllister 2007. Metrics or Peer Review? Evaluating the 2001 UK Research Assessment Exercise in Political Science. Canberra: Australian National University, Research


Copyright of Science & Public Policy (SPP) is the property of Beech Tree Publishing and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder’s express written permission. However, users may print, download, or email articles for individual use.