

# Agenda

Advancing economics in business

## Was it worth it? How to evaluate policy

There is a growing need for robust *ex post* evaluation of policy interventions, the impetus for which is a desire by policy-funders for more effective accountability and decision-making. This article suggests that evaluation is not fraught with difficulties, as is often perceived to be the case, and sets out how a successful evaluation could be designed, building on Oxera's recent work for the UK Department for Transport

### Evaluation and the policy cycle

Evaluation is the process of reviewing the impacts of policy-making. It should be planned from the moment a decision has been made, as impacts are likely to start from this point. *Process* evaluation considers the actions taken by policy-makers and those involved with implementation up to the point at which a project or policy has been completed. *Impact* evaluation assesses whether the impacts of the intervention are in line with the expectations of those taking the original decision. It attempts to determine why outcomes are different from predictions, in order to improve future decision-making. It should not be confused with *appraisal*, which is analysis carried out before an intervention occurs to determine whether it should take place at all.

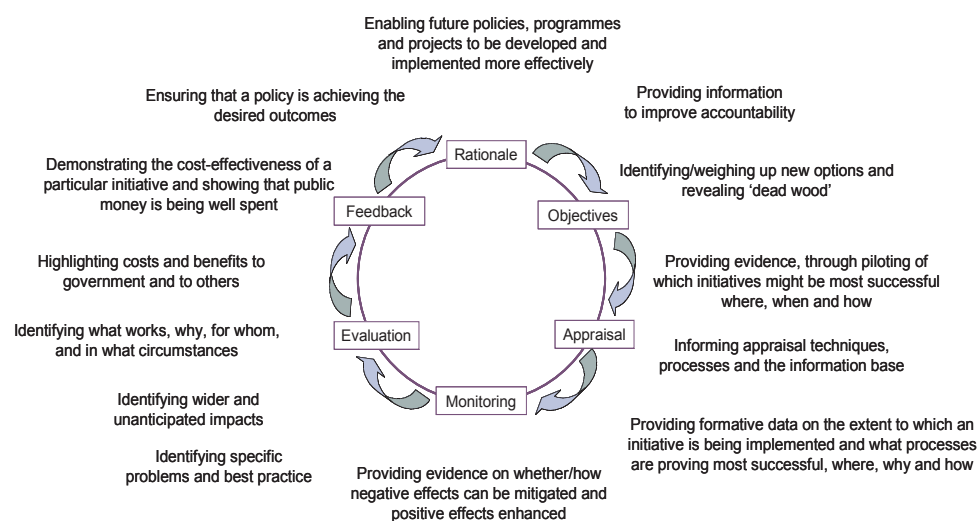
Appraisal and evaluation are both part of the 'policy cycle' (see Figure 1). According to the cycle, policies are

conceived in line with government objectives, based on information revealed by the success or failure of previous interventions. They are then thoroughly appraised, before the policies given the go-ahead are implemented and their impacts monitored. The next stage in the cycle is evaluation, followed, crucially, by dissemination. Without evaluation or dissemination of the evaluation's findings, the policy cycle is broken and the lessons that could be learned from previous interventions are ignored.

### Impetus

The use of evaluation as a key element of the policy cycle has been given fresh impetus in recent times by a number of publications. Most importantly in the UK has been HM Treasury's revised Green Book, which provides binding guidance on government departments in relation to the evaluation of policy interventions.<sup>1</sup> The following is the first paragraph in the evaluation chapter:

Figure 1 The role of evaluation in the policy cycle



Source: Oxera (2005), op. cit.

Oxera's report, 'How Should the Ex Post Evaluation of Trunk Road Schemes be Enhanced?', prepared for the Department for Transport by Oxera, Mott MacDonald, Social Research Associates, Tavistock Institute and University of Oxford Transport Studies Unit, June 2005, is available at [www.oxera.com](http://www.oxera.com).

When any policy, programme or project is completed or has advanced to a pre-determined degree, it should undergo a comprehensive evaluation. Major or ongoing programmes, involving a series of smaller capital projects, must also be subject to ex-post evaluation. (para 7.1, p. 45)

Audit and enforcement of the way government departments undertake their evaluations is, according to the Green Book, to be carried out by the Public Accounts Committee, the National Audit Office (NAO), the Office of Government Commerce's Gateway Process, and departmental arrangements (eg, internal audits and other institutional arrangements to ensure that projects and policies are delivering value for money). An example of this auditing is the NAO's report on the Department for Transport's work in funding the construction of light-rail schemes,<sup>2</sup> which was critical of the Department's incomplete evaluation of existing systems, leading to:

an incomplete picture of what has been delivered for the significant amount of public monies invested in the schemes, and does not have as informed a base as it should have for the consideration of future schemes.

Further impetus has arrived in the form of binding guidance from the Office of the Deputy Prime Minister on assessing the impacts from spatial interventions (typically regeneration, renewal or regional development),<sup>3</sup> the Cabinet Office's 1999 'Modernising Government' White Paper, and the Gateway Process, which comprises a series of appraisal and evaluation 'gateways' that government projects have to pass.

In Europe, large EU interventions are typically associated with substantial evaluations. For example, there has been detailed guidance in place for a number of years (following modest beginnings in 1988) on the evaluation of European Structural Funds.<sup>4</sup> At a national level, evaluation has developed most quickly in countries with strong cultural links with the USA, where evaluation first emerged as a distinct area of professional practice. However, it is of note that the 2003 guidance referred to above observes that evaluation expanded considerably in the UK only with the arrival of a new government in 1997, suggesting that the UK lagged behind other countries in its development of evaluation.

## What are the potential benefits?

Why are funders and policy-makers so keen on evaluation? Oxera's recent study for the UK Department for Transport on the feasibility of enhancing the ex post evaluation of trunk road schemes determined that there are a number of potential benefits from well-designed and disseminated evaluation. Based on information from previous evaluations, a number of consultations, and

in-depth case studies of recent evaluations, the report concluded that key outcomes from improved evaluation could include:

- policy accountability—within (for example) the Department for Transport and the Highways Agency, and externally;
- improving appraisal where it is currently weak, or where there is a large degree of uncertainty about its conclusions;
- closing the appraisal–decision-making–evaluation loop (eg, informing decisions on which road scheme choices are most appropriate);
- developing a cumulative knowledge base;
- enabling continued improvements in project processes and implementation.

There are many types of evaluation 'customer', each gaining different benefits from evaluation. The report noted that appraisers want evidence assisting the development of appraisal tools and techniques; policy-makers want to know that schemes deliver planned benefits and that roads policy is proving effective; designers and implementers want good-practice evidence on delivering roads; and communities want to know that their concerns are effectively addressed. However, these benefits are conditional on evaluation forming an integral part of institutional decision-making, which will often require substantial culture changes within organisations. It is not surprising, therefore, that the Green Book suggests that evaluation requires 'management initiative (sometimes political commitment)' to succeed.

## Potential difficulties

Despite the benefits of evaluation evidence, it is noted above that evaluation only really caught on in the UK with the change in government in 1997. The French government, meanwhile, has recently introduced a new evaluation methodology for motorway and high-speed rail projects, because despite 1982 legislation requiring evaluation of each project above a certain monetary threshold, only three have been undertaken.<sup>5</sup> Furthermore, evaluation spend relative to project spend is often small—for example, in 2004, evaluation represented only 0.1% of the £507m Highways Agency major improvements budget.<sup>6</sup> If evaluation is such a good idea, why has it taken until now for government departments to evaluate their decision-making systematically?

Evaluation is often considered expensive and of limited value. There are a number of plausible reasons for this. First, evaluation might be viewed negatively by those taking decisions, particularly when only carried out to 'find out what's gone wrong'. However, this seems to be

more a question of tone—if evaluation becomes part of everyday life (as suggested by a respondent to one of the consultations carried out by Oxera in the recent work for the Department for Transport), its benefits are far more likely to be realised.

The second barrier to evaluation activity is the availability of continuing resources. The study for the Department for Transport suggested that resourcing evaluation studies per se should not be a barrier to evaluating trunk road schemes. Careful planning and targeting of evaluation spend, and the use of professional resources and data from organisations with an interest in the findings from individual studies, should enable robust conclusions to be drawn at relatively low cost. For larger, innovative or controversial interventions, the Green Book recommends that a proportionate approach is taken:

the thoroughness of an evaluation should depend upon the scale of the impact of a policy, programme or project, and to some extent on the level of public interest. (p. 46)

However, the real barrier seems to be ensuring credible commitment to evaluation, from project conception, through to what might be a number of years after implementation. For example, the researchers evaluating the opening of Sheffield Supertram in the north of England concluded that:

This monitoring exercise was only able to assess the regeneration impacts of the full [South Yorkshire Supertram] system for little more than a year. It may well be that impacts become much more evident over a longer time horizon.<sup>7</sup>

This is indicative of a further potential problem faced by evaluators—providing timely information to policy-makers that is of relevance to the issues of the day. It is quite conceivable that, over the ten-year life of an evaluation study, policy priorities, and methods of implementing policy, change. One approach to addressing this problem could involve the development of a ‘theory of change’. This form of evaluation seeks to understand how policy-makers and other stakeholders in an intervention believe it will lead to desired outcomes—ie, what are the mechanisms by which change is expected to occur? For example, a new tram scheme might be expected to create improved access to employment opportunities in a city centre for those living on the outskirts.

A further step could involve investigating contexts in which decisions are being made. Using the same example, the circumstances in which this intervention is expected to lead to the desired conclusion might be a thriving city centre offering increasing employment

opportunities. With the theory of change in place, evaluators can then play an important role by informing scheme developers of changes in mechanisms and contexts that could potentially harm outcomes. A case in point was the Sheffield Supertram, where the evaluation noted that the city centre experienced consistent decline during the scheme’s implementation, whereas the scheme appraisal expected it to be the dominant employment focus. The evaluation concluded that new links to employment areas replacing the city centre as the focus of new employment opportunities, and better integration between regeneration and planning policies are required. Providing this sort of advice during the implementation process would enhance the rationale for evaluation, as it would improve the implementation of the intervention in question, and other similar interventions being undertaken.

The expectation in the Green Book is that the counterfactual (for example, what might have happened in the absence of the intervention) should be developed by evaluators. This is also required by the new French transport evaluation methodology. However, Oxera’s conclusion in its report for the Department for Transport was that developing a counterfactual—which often requires an expensive investigation into appraisal documentation that is not necessarily available—is not required in all cases. Instead, valuable evaluation evidence could be collected using less expensive, more pragmatic approaches, with the counterfactual being developed only when it is proportionate to do so. Where a counterfactual is appropriate, theory-of-change approaches appear to offer a robust solution, with the potential for developing a series of counterfactual positions.

One final issue is how boundaries for impacts can be defined. Who does an intervention affect? In the case of Manchester’s trams, for example, it is arguable that the effects have been much wider than within a five-minute walk of tram stops—they have become part of the city’s image, being used in television clips of the city by national and international media. Perhaps even more difficult to distinguish are the impacts of policies themselves. Again, this is where a theory-of-change approach could help by following through from the outset what the most likely outcomes are, and focusing evaluation effort on the types of people expected to be affected. With a well-specified logic model,<sup>8</sup> it would also be possible to pick up confounding factors or unintended impacts beyond those that the policy was expected to affect.

The box below defines the requirements for a successful evaluation process.

## A recipe for success

The recipe for learning through robust evaluation includes the following ingredients.

- *Early planning*—at a minimum, early planning of evaluations can enable evaluators to determine the main issues associated with the intervention. For larger-scale evaluations, early involvement can lead to the development of an underlying logic model, which can be used for the continuous monitoring of the policy or project. This should lead to outcomes closer to the underlying objectives.
- *Stakeholder/customer involvement*—involving an evaluation's customers in its development and execution seems to be a critical success factor. This can ensure that it appeals to the different types of customer it is being carried out for—from those at a national level seeking to enhance a knowledge base and increase certainty in decision-making, to those directly affected by the intervention at a local level. It can also reduce direct costs—for example, if those being consulted are able to provide access to information being gathered for other purposes. However, it is important that stakeholder involvement does not lead to 'capture' of the evaluators—evaluations should be carried out independently of those who commissioned the intervention, and findings should not be unduly influenced by interest groups.
- *Proportionality*—while potentially beneficial, evaluation is not an end in itself, and evaluation activity should be justified carefully. Oxera's paper for the Department for Transport argued that there should not be a bare minimum of evaluation that should be undertaken by the Highways Agency for each scheme—rather, evaluation activity in respect of each new road should be selected on the basis of

the needs of local and national users of the evaluation's outputs.

- *Effective dissemination*—without dissemination, the value of evaluation is limited. A key finding of the study for the Department for Transport was that:

Many consultees were unaware of evaluations or of their findings. It is not surprising that there is unmet demand for evaluation evidence if its customers are unaware that it is being conducted and have no opportunity to shape it into something that meets their needs. Likewise, it is not surprising that some stakeholders cannot immediately appreciate the potential benefits of evaluation if they are unaware of its existence.

- *Tracking and using a cumulative knowledge base*—dissemination should be augmented by a record of evaluation findings, which grows over time to provide a searchable knowledge database covering all aspects of the types of intervention being evaluated. This will ensure that not only today's evaluation customers, but also tomorrow's, are able to benefit. It also enables what is termed 'meta evaluation', whereby analysis is undertaken using the findings from a number of similar evaluations. The Highways Agency recently used this technique successfully to assess which types of small road scheme provide the most beneficial outcome.

Following this relatively simple recipe should lead to robust evaluation evidence, focused on answering pertinent questions in a way that appeals to a diverse set of customers. Its ultimate goal is to tell us what works, for whom, and in what circumstances.

<sup>1</sup> HM Treasury (2003), 'The Green Book: Appraisal and Evaluation in Central Government—Treasury Guidance'.

<sup>2</sup> NAO (2004), 'Improving Public Transport through Light Rail', report by the Comptroller and Auditor General, HC518, Session 2003–04, April 23rd.

<sup>3</sup> Office of the Deputy Prime Minister (2004), 'Assessing the Impacts of Spatial Interventions: Regeneration, Renewal and Regional Development—The 3Rs Guidance'.

<sup>4</sup> The latest version is: Tavistock Institute (2003), 'The Evaluation of Socio-economic Development: The GUIDE', December, prepared in association with GHK and IRS. This replaces the *MEANS Handbook* to evaluating European structural funds interventions.

<sup>5</sup> Chapulut, J.N., Taroux, J.P. and Mange, E. (2005), 'The New Ex Post Evaluation Methods for Large Projects in France', proceedings of the European Transport Conference, Strasbourg, October 3rd–5th.

<sup>6</sup> Source: Highways Agency Business Plan, 2003/04.

<sup>7</sup> Centre for Regional Economic and Social Research (2000), 'Monitoring the Economic and Development Impact of South Yorkshire Supertram', Sheffield Hallam University, para 10.7.

<sup>8</sup> A logic model is the practical manifestation of a theory of change. It is a computer-based model of contexts, mechanisms and outcomes, including common linkages, associated with the intervention in question.

If you have any questions regarding the issues raised in this article, please contact the editor, Derek Holt: tel +44 (0) 1865 253 000 or email [d.holt@oxera.com](mailto:d.holt@oxera.com)

Other articles in the October issue of *Agenda* include:

- competition as a public policy tool: what is the evidence? *Andrew Rees and Sasha Maguire, DTI*
- reinsurance in the EU: voluntary or mandatory regulation?
- public information, private profit: how should government agencies compete?

For details of how to subscribe to *Agenda*, please email [agenda@oxera.com](mailto:agenda@oxera.com), or visit our website

[www.oxera.com](http://www.oxera.com)