

Evaluation and its use in value for money studies

Learning from Others

*Our doubts are traitors,
And make us lose the good we oft might win
By fearing to attempt*

Measure for Measure Act I, Scene iv



*There is a tide in the affairs of men,
Which, taken at the flood, leads on to fortune;
Omitted, all the voyage of their life
Is bound in shallows and in miseries*

Julius Caesar Act IV, Scene iii

Evaluation and its use
in value for money studies

Learning
from Others

*Our doubts are traitors,
And make us lose the good we oft might win
By fearing to attempt*

Measure for Measure Act I, Scene iv

Foreword

Change is a key feature of all aspects of the work which we do in the National Audit Office. This is brought very much into focus as the Government takes forward its modernising programme to improve all aspects of service delivery for the benefit of citizens. New ways of delivering government programmes are being developed and seemingly intractable social and economic problems – “wicked” issues are being looked at afresh. To report objectively and constructively on how these new initiatives are being implemented and to make recommendations which are forward looking requires us to modernise as well. While we have a good track record of producing high quality Value for Money (VFM) reports we need to continue to enhance all aspects of our work. We can do this by enhancing our skills, by learning from others, by being receptive to new ideas, and by developing our knowledge pool that all colleagues can share.

This guide can help you do this by drawing on the approach adopted by the evaluation community who often carry out studies similar in many respects to VFM examinations using a wide range of diagnostic and analytical approaches.

For further advice on any aspect, methodology or technique discussed in this guide, or help in arranging training, please contact the VFM Development Team – telephone 020 7798 7175

Good counsellors lack no clients

Measure for Measure Act I, Scene ii

Contents

Foreword	3
Introduction	5
1 Different types of evaluation and how they can help in VFM study design	17
2 Key stages in an evaluation and some helpful concepts	41
3 Evaluative techniques and how they can help in VFM examinations	61
Bibliography	88
List of Credits	90



Introduction

*And as the morning steals upon the night, Melting the
darkness, so their rising senses Begin to chase the ignorant fumes
that mantle their clearer reason*

The Tempest Act V, Scene I

What is evaluation?

- 1 The term evaluation is widely used in many contexts to cover a range of different judgements. For example, if you attend the theatre or see a film you form a personal opinion of whether you liked or enjoyed what you saw; if you go to a football match you form a view of the skill of the players. These are very much informal assessments which we make everyday in all aspects of our life. Evaluations can, however, be much more rigorous and formal involving detailed assessments of the achievement of for example, health outcomes such as care for the elderly or programmes to reduce social exclusion, to prevent juvenile crime or to reduce hospital acquired infection. **Formal evaluation** is a disciplined inquiry that applies scientific procedures to the collection and analysis of information about the content, structure and outcomes of programmes, projects and planned interventions.¹
- 2 Evaluation as a discipline does not have a methodology of its own. It uses a range of diagnostic and analytical methods such as questionnaires, surveys, observation, interviewing and data analysis often drawn from the social sciences. Evaluation relies on both quantitative and qualitative data and is concerned with determining the merit, worth or value of an established policy, programme or planned intervention.

Is evaluation different from value for money work?

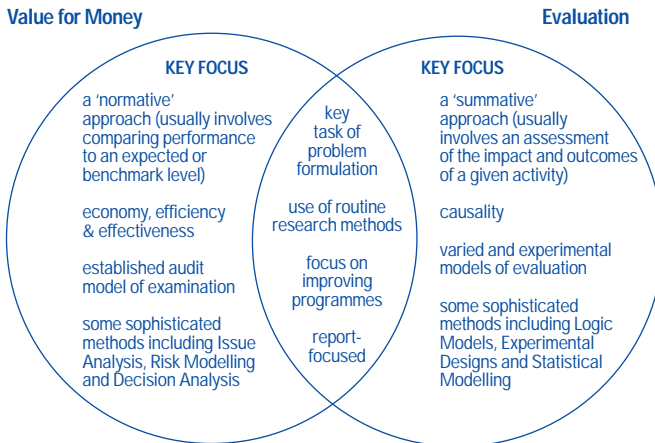
- 3 The reason why evaluation can be such a useful discipline from which to learn is that it has grown up in an entirely different (non-audit) environment from the value for money discipline. Many of the fundamental principles of formal evaluation were developed in the United States in the late 1960's and early 1970's, drawing heavily from

¹ Evaluation Research Alan Clarke ISBN 0 7619 5095 8 1999.

the fields of Social Science and public policy analysis. This background has led to a similar focus to value for money work (on government programmes and interventions) but with a much greater emphasis on:

- Outcomes
 - Causal relationships
 - The views of all stakeholders at all stages, and
 - Sustainable recommendations
- 4 This in turn has influenced the models of evaluation used (illustrated on page 15) and the methodologies routinely employed (see Part 3). Between the two forms of examination – evaluation and value for money – there are similarities, but there are also differences. These are further illustrated below and in the Figures overleaf.

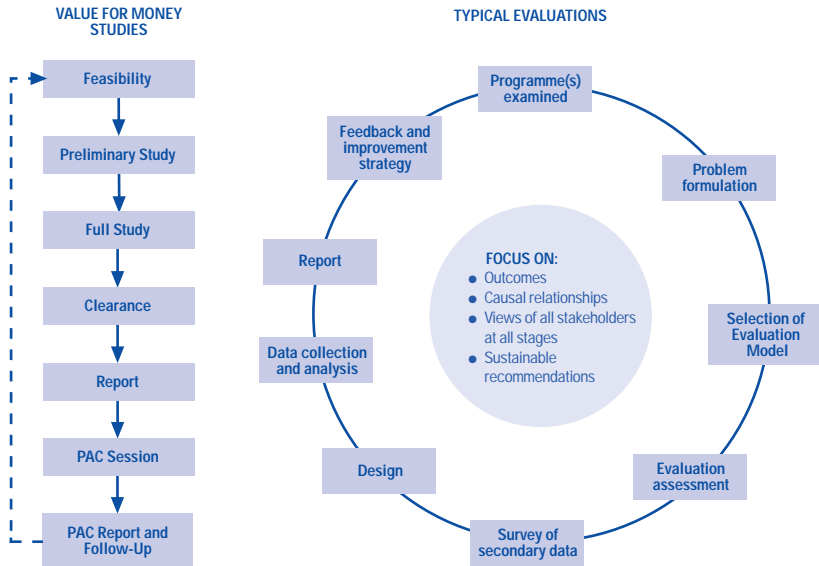
Diagram showing the overlap and differences between the Value for Money and Evaluation disciplines



Differences of emphasis between evaluation and value for money examinations (overleaf)

Features	Value for money Examinations	Evaluation
1. Mandate	The mandate for the NAO to undertake VFM examinations of central government is set out in legislation. Ideas for VFM examinations are usually developed by the NAO.	Evaluations are usually undertaken at the request of programme sponsors. Some may be undertaken by external evaluators e.g. General Accounting Office.
2. Scope	VFM examinations are typically 'normative' in that programme performance is measured against policy objectives. And policy objectives can not be questioned as part of a VFM examination.	Evaluations are characteristically 'summative' in that performance is examined in terms of outcomes arising from a programme(s). Policy objectives may be questioned as a part of the evaluation.
3. Models of examination	The normative model is an important influence on how VFM examinations are conducted, although by no means the only model adopted (for example, VFM examinations also make use of economic models amongst others).	Evaluation tends to draw on a wider range of different models. For example, 'Goal-free' evaluations involving deliberately setting aside the normative goals of a programme in favour of an examination of observable outcomes. Other evaluations may be founded in the perceptions and views of programme stakeholders.
4. Focus	VFM examinations typically focus on economy, efficiency and effectiveness with a view to securing financial savings and qualitative improvements to government programmes.	Evaluations usually focus on cause and effect relationships, with a view to determining how programme effects are produced and how changes to particular elements of the programme will affect outcomes.
5. Methods	A wide range of methods are used. Typically VFM examinations have relied upon file examinations, interviews, surveys and costing exercises to obtain evidence. More recently the VFM discipline has made advances towards more innovative techniques such as modelling, statistical analysis and software-based techniques .	A wide range of methods are used, including interviews and participatory techniques, Experimental Designs, Logic Models, Meta-Analysis and advanced statistics .
6. Output	VFM examinations normally lead to a published report. The VFM team does not usually become involved in implementing recommendations.	Evaluations do not usually lead to a published report. Typically a report is made available to the sponsoring body (and possibly stakeholder groups). The evaluation team may be asked to assist in implementing recommendations.

Key stages in a typical evaluation compared to VFM studies



Why is evaluation important for value for money examinations?

- 5 The discipline of evaluation can provide a useful **framework for thinking** about how we can approach value for money studies. In the early stages of all VFM examinations we have to make choices about how we should frame and diagnose issues, which methods we should use; how we should ensure that we have the right data to answer our questions and how to deliver engaging reports. Evaluators face the same questions and can offer a fresh perspective for VFM work.
- 6 Furthermore, having some understanding of evaluation is important and may be helpful in providing insights into:

- i **Assessment of outcomes.** Evaluations can provide evidence on outcomes in various ways. **Firstly**, by addressing basic questions such as why a certain outcome exists for example high levels of juvenile drug abuse. This type of evaluation often known as **ex ante** might focus on cause and effect with the aim of recommending a series of interventions – better education, more targeted policing to improve policy implementation so that drug abuse is reduced. **Secondly**, monitoring of longer term programmes might include provision for a series of formal evaluations at key stages to provide reliable data on the impact of the programme and how it might need to be improved. **Thirdly**, at the end of a project or programme evaluations can be important in assessing their success in achieving their objectives and securing planned outcomes, their likely sustainability, and lessons for the future.

*Within a continuous cycle of policy development, evaluation is a key input to the appraisal of new policies. Evaluation of past policies enables the design and delivery of current and future policies to be improved in light of experience – **Adding it Up – Improving analysis and modelling in central government – A Performance and Innovation Unit Report – Cabinet Office January 2000.***

- ii **Improved policy making.** The Modernising Government White Paper (Cm 4310) published in March 1999 set out a major long term programme to improve all aspects of government so that tax payers get a better deal for their money. One aspect of this is improving policy making so that programmes are designed to ensure that they deliver planned outputs and outcomes and are responsive to citizens' needs. To help achieve this the White Paper recognised the need for evaluation to play a greater part in designing programmes so that policy making is much more evidence based.

In our VFM work we may need to form a judgement as to whether departments and agencies are commissioning evaluations at the most appropriate time, that they are well designed and of the right quality.

- iii **Examination of “wicked” and cross-cutting issues.** One aspect of the Modernising Government agenda is to give fresh impetus to tackling seemingly intractable problems such as long term employment or social deprivation. One response is to develop programmes in a more joined up way so that all possible government interventions are co-ordinated to ensure that they are complementary, and that by working more in unison departments and agencies develop new innovative approaches. As different approaches are being developed they are often piloted involving partnerships between local and central government, the private, public and voluntary sectors. Formal evaluations of the success of these pilots and how they might be improved before they are rolled out nationally are becoming a key component of policy development.

Joining up – what it means:

“ The government’s focus on cross-cutting outcomes presents a major challenge to policy makers. The Modernising Government White Paper envisages policy making as built around shared goals, not around organisational structures or existing functions. But joining up is not just about shared approaches to cross-cutting issues. Horizontal joining up between organisations needs to be supplemented by better co-ordination between policy makers in the same departments and by better ‘vertical’ joining up with service deliverers and those who implement policy. It is not an end in itself but should be undertaken where it adds value.” **Professional Policy Making for the Twenty-first Century Cabinet Office (September 1999).**

As our VFM work focuses more on the success of programmes to tackle “Wicked” issues we will need to draw on the evaluations of pilot schemes and assess how the lessons learned from them are being implemented.

- iv **Evaluations** as a source of evidence. As more prominence is given to evaluations by departments and agencies in developing policies and reviewing their effectiveness evaluations will become an important source of evidence for our VFM examinations.

Using evaluations as a source of evidence requires specialist skills to interpret their conclusions and overall validity. We will need to make greater use of techniques such as meta analysis (explained in part 3) to synthesise the results of different evaluations to reach conclusions which are authoritative.

How this guide can help

- 4 This guide provides a basic introduction to the principles of evaluation and how they can help improve the quality of our VFM examinations. The guide is structured so that:

Part one	explains the different types of evaluations and the circumstances when they are most appropriate.
Part two	sets out the key stages in an evaluation and some helpful concepts.
Part three	discusses some of the techniques used in evaluations to collect and analyse data and their relevance to VFM examinations.

A bibliography for further reading and reference is provided on page 88.

The NAO is a member of

The American Evaluation Association

(Web site <http://www.eval.org/>)

European Evaluation Society

(Web site <http://www.europeanevaluation.org/>) and

UK Evaluation Society

(Web site <http://www.evaluation.org.uk/>).

Other useful sources are the

Australasian Evaluation Society:

(Web site <http://www.parklane.com.au/evalnet/>) and

The Canadian Evaluation Society:

(Web site <http://www.evaluatincanada.ca/>).

Anyone requiring further advice should contact

Michael Whitehouse

(telephone 020 7798 7078

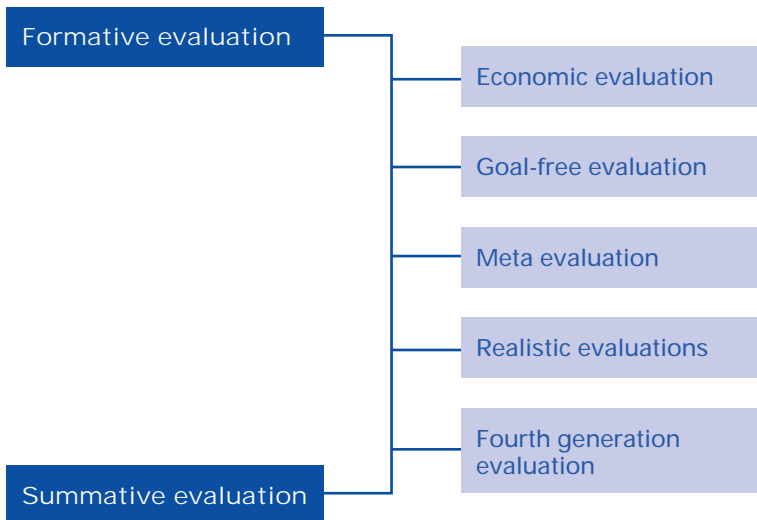
e-mail: michael.whitehouse@nao.gsi.gov.uk) or

David Goldsworthy

(telephone 020 7798 7514

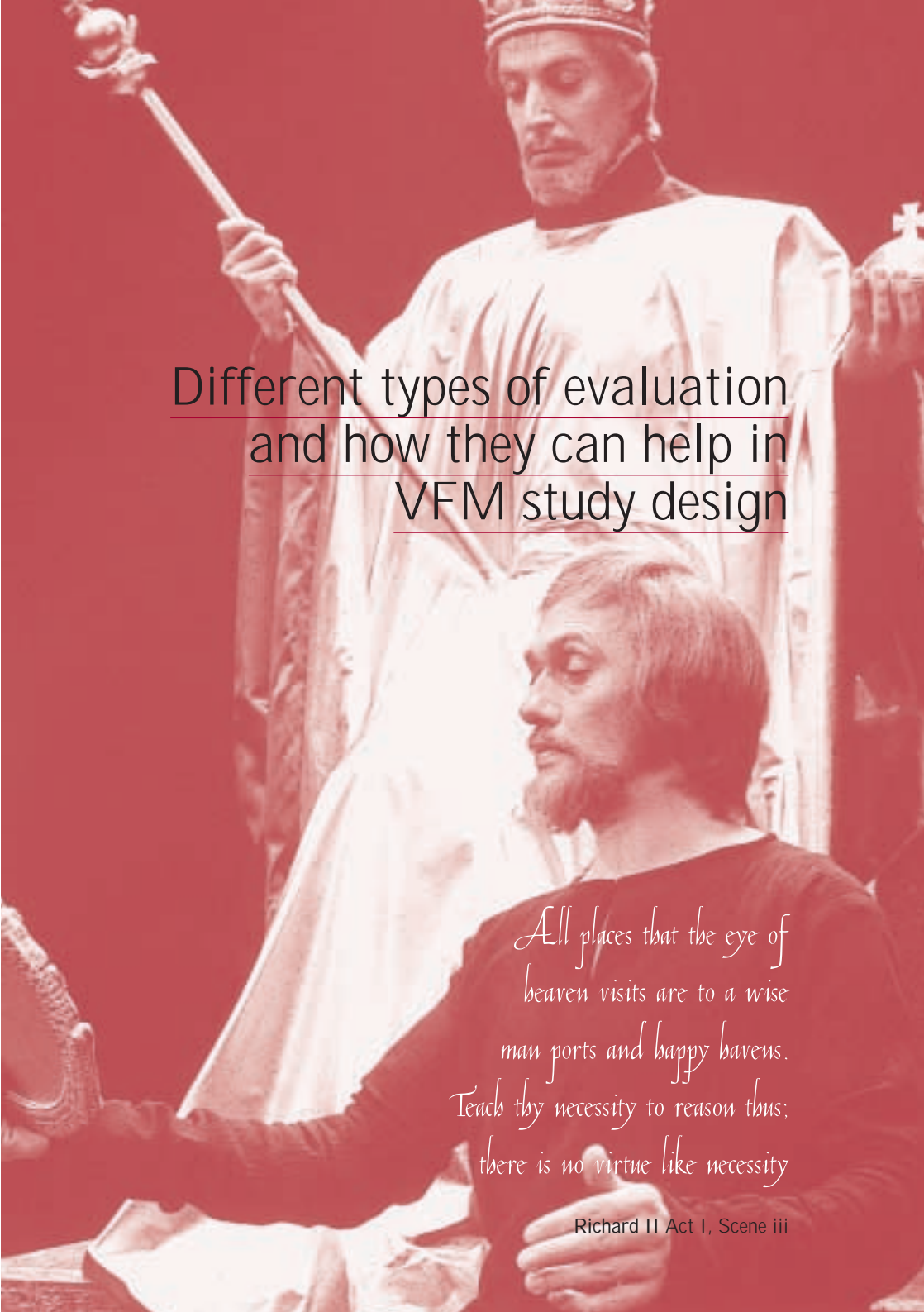
e-mail: david.goldsworthy@nao.gsi.gov.uk)

Different types of evaluation



Be sure of it; give me the ocular proof

Othello Act III, Scene iii



Different types of evaluation
and how they can help in
VFM study design

*All places that the eye of
heaven visits are to a wise
man ports and happy havens.
Teach thy necessity to reason thus:
there is no virtue like necessity*

Richard II Act I, Scene iii

- 1.1 Evaluation is essentially about critically assessing the knowledge you have gained about a programme, project or policy to form an objective judgement about its success or failure. It can take various forms with different shades of emphasis and focus. These different foci are often reflected in the types of value for money studies which we do.

Knowing about different types of evaluations and the sorts of questions which they address can help you in designing VFM studies. They can help you think laterally to ensure that your studies get to the heart of the problem and come up with recommendations which add value, are practicable, and which will deliver long term sustainable benefits.

- 1.2 This part of the guide tells you about (i) the different types of evaluation and (ii) evaluations as a part of the policy making process.

(i) Different types of evaluation

- 1.3 There are basically two different types of evaluations:
- a **Formative evaluations** are involved with identifying the strengths and weaknesses of a programme or intervention. Information is collected on the design and implementation of the programme or project to form an assessment of their overall effectiveness. The aim is to ascertain if any changes are needed in order to improve the programme. Formative evaluations are therefore ongoing and will be carried out periodically throughout the life of a programme. *A good example of a VFM study which was **formative** in its focus was our study of the Management and Control of Hospital Acquired Infection in Acute NHS Trusts in England (HC 230 1999–00).*

Formative Evaluation

The Management and Control of Hospital Acquired Infection in Acute NHS Trusts in England (HC 230 1999 – 00)

This report looked at arrangements within National Health Service hospitals to minimise infections that patients can acquire while they are in hospital. Each year around 100,000 patients acquire infections, leading to longer periods of hospitalisation, additional costs and, in some cases, medical complications. The study examined strategic management arrangements within hospitals and the measures taken to detect, monitor, prevent and control acquired infections. The report made recommendations to the National Health Service Executive, Health Authorities, and hospitals. These included better sharing of management information, the integration of infection control arrangements with the bed management system and a review of hospital hygiene. An important focus of the report was on how hospitals can improve their own arrangements rather than identifying the impact or wider effects of the prevention and treatment programme.

1.4 The typical questions which a formative evaluation or study might ask are:

- 1 Are the aims and objectives of the programme clearly formulated? And is there common understanding of their meaning and commitment to their achievement by all those involved in the implementation of the objectives?
- 2 Is the process of implementing the programme cost effective?
- 3 Was the programme implemented as intended?
- 4 Is there any evidence of under performance requiring programme implementation to be redesigned or remedial action taken?

- 5 If underperformance exists what are the key factors responsible for the underperformance and how should these be dealt with?
 - 6 Are those intended to benefit from the programme or activity or most influenced by it satisfied with the service they are receiving?
 - 7 Is the programme or policy being implemented in a sufficiently joined-up way so that it contributes to the Government's wider policy objectives?
- b **Summative evaluations** are about forming a view about the overall effectiveness or impact of a programme or project. In government summative evaluations are often designed to provide policy makers with independent information on whether a programme should continue to run or be fundamentally changed. The emphasis is very much on the ultimate outcomes of a programme. Given the Modernising Government programme's emphasis on outcome achievement we need to consider how we can identify these outcomes and measure them in our VFM studies. *A good example of a VFM study which was summative in its focus was our study of Cataract Surgery in Scotland (HC 275 1997 – 98).*

Summative Evaluation

Cataract Surgery in Scotland (HC 275 1997 – 98)

This report examined the arrangements within the Scottish Office Department of Health for undertaking cataract operations. Cataract surgery is one of the commonest procedures carried out in the NHS, with most patients reporting an immediate and dramatic improvement in sight after only a 20 minute operation. The study that was undertaken focused on the outcomes associated with day surgery as opposed to in-patient, overnight surgery – which was, at the time, the approach adopted in the majority of cases. The investigation mapped the costs and

benefits of the two forms of cataract surgery, giving particular emphasis to the impact on patients. The study team reviewed patient surveys commissioned by the health service, independent surveys and published papers in order to form a view on how increased levels of day surgery might affect patients. The study also considered the impact on patient waiting times, the clinical efficiency of day surgery and the information needs of the patient. As a result the study team were able to recommend day surgery as an approach that maximised the benefits to the patient with wider outcomes in terms of successfully treating more patients at reduced cost.

1.5 The typical questions which a summative evaluation or study might ask are:

- 1 Has the programme or policy achieved its intended outputs and outcomes?
- 2 Are these outputs and outcomes likely to be sustainable over the time period originally intended when the programme was designed?
- 3 Are those intended to benefit by the programme, project or any other intervention actually benefiting and is there any unintended social or geographical exclusion?
- 4 Are there any negative side effects which the output or outcome has contributed to which should have been avoided or for which some remedial action is needed?
- 5 Is the programme or project supportive of or complementary to wider policy initiatives?
- 6 What are the key lessons which are likely to improve the achievement of outputs and outcomes in future programmes?

The essential difference between the two types of evaluations is that “ summative evaluations tend to be conclusion-orientated whereas formative evaluations tend to be action-orientated” ²

1.6 Within these two general types of evaluation there are five subsets which we partly reflect in the types of VFM studies which we do. They are:

- i **Economic evaluations** are concerned with assessing likely future costs, revenue streams and economic benefits such as job creation, quantifying risks and wider economic benefits. Economic evaluations will usually involve greater quantification using techniques such as net present value, discounting, Monte Carlo simulation and Bayesian modelling to predict probabilities. Our studies of Private Finance Initiative projects and public private partnerships in general, such as our report on the contract to complete and operate the A74(M) M74 Motorway in Scotland (HC 356 1998 – 99), are good examples where we seek to form an economic as well as a financial view as to whether the best deal was obtained. As we do more studies to assess the success of introducing more commercialisation into the public sector economic evaluations are likely to become more important.

² M. Q. Patton (1986) *Utilization – Focused Evaluation*, 2nd edn, Newbury Park, CA, Sage.

Economic Evaluation

The Private Finance Initiative: The Contract to Complete and Operate the A74(M) M74 Motorway in Scotland

For a privately financed road to represent value for money, the price to be paid must be in line with the market, the contract must provide a suitable framework for delivering the service specified and the cost of the privately financed option, taking account of risk, should be no more than that of a publicly funded alternative. We examined this contract to determine whether these criteria were met. We found that to assess whether the benefits of a privately financed road were likely to outweigh the extra costs of private finance, the Department constructed a public sector comparator and reviewed the bid against a number of sensitivities including a range of discount rates.

As well as capital costs the public sector comparator for the contract needed to take account for differences in operating and maintenance costs between public and private solutions over the 30 year comparison period. The Department's analysis showed that the privately financed road in Scotland should cost £17 million less than the publicly funded alternative. Part of the justification for a three lane motorway included a cost benefit analysis of the predicted savings in road users' time and from reduced accident rates both during normal operating conditions and during periods of future maintenance by virtue of a reduction in traffic congestion.

Economic evaluation – Some terms

Note: strictly, the term "economic evaluation" is used in government to refer to ex-post analysis of projects. The usual term for *ex-ante* analysis is "appraisal".

Discounting. The process of converting quantities, which are distributed over time, to a “present value” (see below). This is done by reducing the quantities over time to take account of the rate at which the value of money (or other unit of account for non-financial quantities) declines through time. This rate is known as the “Discount rate”.

Present Value (PV). The sum of discounted values of a future stream of costs (Present Value of Costs, PVC) or benefits (Present Value of Benefits, PVB).

Net Present Value (NPV). The difference between the PVC and PVB. If this is positive, then the project’s benefits are greater than its costs. This can be used as a decision criterion; for example, if the NPV is positive, then the project will go ahead¹.

Internal Rate of Return (IRR). This defines the discount rate at which the NPV of a product is zero. This rate can be compared with; for example, the government’s discount rate of 6% real. This can be used as a decision criterion, such that a project will go ahead if the IRR is greater than a set target rate of return¹.

Opportunity Costs. The value of opportunities foregone because of an intervention.

Shadow prices. The imputed or estimated costs of goods and services when these goods and services are not valued in the current market place.

Benefit-Cost Ratio. The ratio of the Present Value of Benefits to the Present Value of Costs. If the ratio is greater than one, then the project’s benefits are greater than its costs. Target ratios can be set as a decision criterion; for example, if a project achieves a ratio of 1.3:1, then it will go ahead¹.

Monte Carlo Simulation. A method used mainly in Risk Analysis to estimate the likely impacts of risks/uncertainties on key project outputs, such as costs. Spreadsheet based computer packages enable this form of analysis to be undertaken fairly easily.

For example, the Highways Agency's analysis of the value for money of the first four Private Finance Design, Build, Finance and Operate roads included Monte Carlo simulation based estimates of the likely costs of construction and operation risks. These were estimated using such information as engineering data on ground conditions and experience with other projects. The Agency used the calculations to estimate monetary values of the amount of construction and operation risk which was transferred to the private sector consortia under the private finance deal. They used these estimates to judge the value for money of the private sector bids compared with a public sector comparator. If the Present Value of Costs of the private sector bids was lower than that of the public sector comparator, including transferred risk estimates, then the deal was judged to be value for money.

The comparison of the private sector bids with the public sector comparator was discussed in our report: "The Private Finance Initiative: The First Four Design, Build, Finance and Operate Roads Contracts" (HC 476, Session 1997 – 98).

Statistical Decision Processes. These offer a structured approach to choosing between a number of alternatives in a situation of uncertainty. They require the definition of a utility function (for example, profit) whose value depends upon which alternative is chosen, and on other variables subject to uncertainty (for example, sales). A probability distribution has to be attached to each such variable. For each possible alternative, a decision criterion, such as the average value of the utility function across all realistic scenarios, is examined, and the

alternative chosen is the one to optimise (typically maximise) this criterion.

Bayesian decision processes. These are decision processes in which the probability distributions attached to relevant variables are updated as new information becomes available. The best alternative to use may thus change over time.

Note: 1. All targets for decision criteria stated here are illustrative examples only. Parties undertaking the analysis for a particular project will decide which criteria and target figures are appropriate for their project. For further advice contact Kevin Browne telephone 020 7798 7753 e-mail: kevin.browne@nao.gsi.gov.uk

- ii **Goal free evaluations** do not start with the specific objectives of a programme but with the client group most affected by the programme. The aim is to determine the influence of the programme by ascertaining how the client group is affected and comparing this with what the programme is specifically targeted to achieve. In this way the evaluation is not at risk of inadvertent bias because the evaluator is only aware of the client group and that a programme exists but not the programme's precise objectives. A good example would be evaluating programmes to reduce obesity in a certain sector of the population. You could initially examine the trends in obesity and the causes behind it independently of the government's specific interventions to reduce it. You would then seek to correlate your assessment of the underlying causes back to the actions government was taking to assess their appropriateness and potential for success. The goal free evaluation approach takes time and is likely to be more expensive but may be suitable for assessing the success of departments initiatives to promote more joined up government for example, by focusing on the quality of service received by those intended to most benefit most from a joined up initiative.

Goal-Free Evaluation

HM Prison Service: Prison Catering (HC 277 1997 – 98)

Our study of prison catering looked at two main aspects of prison catering: the quality of food provided and the costs involved. With regard to quality standards, the Prison Service did not have performance targets or explicit objectives. Higher level objectives such as the need to provide decent conditions for prisoners and meet their needs were the only performance benchmarks available. The study team therefore adopted a goal-free approach and commissioned consultants to visit a range of prisons to consider the issues presented. Through discussions with prisoners and surveys the team built a picture of what was important to prisoners. The approach focused efforts on the presentation of food, the choices on offer, the balance of food types, the size of portions and the quality of the meals. Standards were compared between prisons and related back to the processes put in place at each establishment. The study recommended that prisons should more systematically compare the standards of catering they provide and develop self-assessment standards to apply to their catering arrangements.

- iii **Meta evaluations.** Often a programme or activity may be subject to a series of evaluations at different stages in its design and implementation. The evaluations may also be undertaken from a range of different perspectives – social, economic as well as the views of the stakeholders most affected by the programme. These evaluations may or may not have influenced the future course of the programme and its implementation but we may need to know about this because it may be a significant factor in the success of the programme. In this sense evaluations become a source of evidence. But, we cannot take them at face value we need to examine them and form a view of their quality and the context

within which they were carried out. The process of doing this is commonly known as “meta evaluation” and as a technique this is discussed in more detail in part three.

- iv **Realistic evaluations.**³ These types of evaluation are less concerned with the theory behind why a programme’s impact was achieved in a certain way – for example assessing the strength of the linkages between causes and effects. They are more concerned with

“When the cook tastes the soup, it is formative evaluation; when the dinner guest tastes the soup, it is summative evaluation”⁴

looking at the specific actions of individuals and groups of people – the social dimension – and how these are contributing to the success of a programme. A good example of this is our study *Giving Customers a Choice The Introduction of Competition into the Domestic Gas Market (HC 403 1998 – 99)*. In this report we were concerned with the success of the opening up the market to competition. Our assessment of the competitive model adopted was very much influenced by the actions and views of the consumers affected by it.

Realistic Evaluation

Giving Customers a Choice – The Introduction of Competition into the Domestic Gas Market (HC 403 1998 – 99)

Following the introduction of competition domestic gas users can now choose which company they buy their gas from. Our study examined the effectiveness of competition not from the perspective of the way competition was introduced but from the benefits which customers were receiving and their

³ R. Pawson & N. Tilley (1997) *Realistic Evaluation*, London, Sage.

⁴ *Evaluation cookbook*, The learning Technology Dissemination Initiative Heriot-Watt University, Edinburgh – ISBN 09528731 68.

perceptions of the benefits. We therefore examined whether the price which customers were paying for their gas had reduced and whether all were benefiting to the same extent; whether they had more choice of supplier and whether this choice was likely to be sustainable in the longer term; whether customers were receiving a better quality of service; and whether gas safety was guaranteed. A major source of our evidence was a nation wide survey of customers which we commissioned MORI to carry out.

- v **Fourth generation evaluations** focus on the claims, concerns and issues of stakeholders. The starting point for this type of evaluation is to capture and understand the perceptions of those groups involved with and affected by the programme examined. This marks a departure from the 'normative' value for money tradition of comparing the actions taken within a programme to stated goals or objectives. Instead programme objectives, the associated risks and benefits are all seen 'through the eyes' of stakeholders. An example that brings out some of the features of Fourth Generation Evaluation is our study *Major Equipment Storage (HC 1006 1997 – 98)*. In this report we identified thirteen key players involved in making decisions about equipment storage requirements. This approach helped to pinpoint where individual objectives fitted with one another and where gaps and misapprehensions lay.

Fourth Generation Evaluation

Ministry of Defence: Major Equipment Storage (HC 1005 1997 – 98)

The Major Equipment Storage report focused on the Ministry of Defence's equipment storage task which included 204 aircraft, 9,166 ground equipments and 19 Naval vessels. In total there were 15 separate Defence agencies and divisions within the Ministry of Defence dealing with different aspects of equipment storage from strategic planning down to technical implementation. As a result of this there were many stakeholders within the organisation each holding different perspectives on the objectives and priorities of the storage task. The study made use of a number of qualitative techniques, such as Cognitive Mapping (explained in Part 3), to draw out the views that were held and to consider how they fitted together. This approach helped to identify common difficulties between the key players, in particular the disincentive effect that budgetary arrangements had on making cost-effective storage decisions. The evaluation approach adopted also helped to locate key weaknesses in management information, based on the needs of stakeholders rather than an alternative model of management information that the VFM team might think is appropriate.

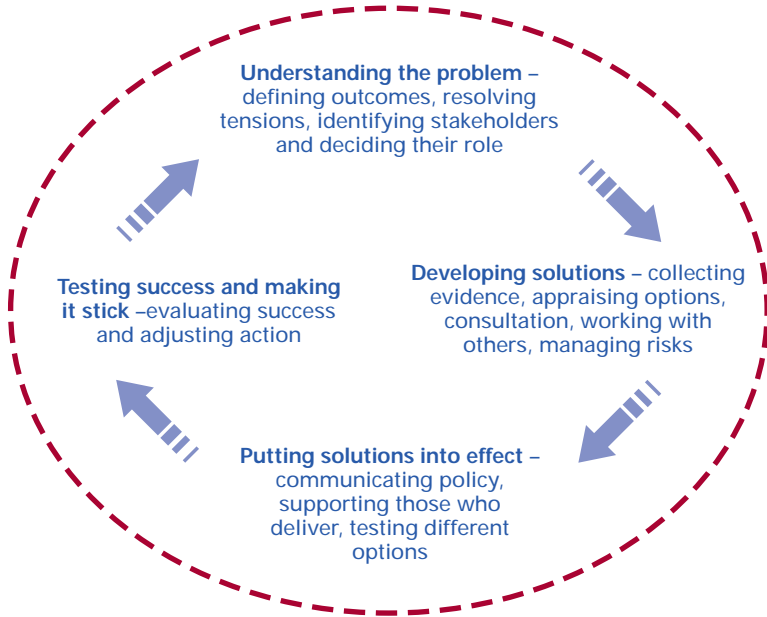
Summary: Although evaluations are similar in many respects to value for money studies, they can take a wider range of different forms depending on the purpose which they are intended to be used for, which in turn will influence how they are designed and implemented. Some knowledge of the evaluator's approach can help us to think laterally to be more innovative in our design of studies and ultimately to improve the value added which our studies provide. The key message is always allow some time to "think out of the box" in designing your study to ensure that you have been sufficiently open to all ideas.

(ii) Evaluation as part of the policy making process

- 1.7 Although our statutory remit prevents us from questioning the merits of policy objectives we need to know the process by which policy is formulated and implemented. This understanding can form the basis of stronger and more informed value for money investigations that get to the heart of policy objectives.
- 1.8 In September 1999 a report – Professional Policy Making for the Twenty-first Century by the Strategic Policy Making Team in the Cabinet Office set out a model for "modernised" policy making. This defined the characteristics of "modernised" policy to be:
- **Strategic** – looks ahead and contributes to long term government goals
 - **Outcomes focused** – aims to deliver desired changes in the real world
 - **Joined up (if necessary)** – work across organisational boundaries
 - **Inclusive** – is fair and takes account of the interests of all
 - **Flexible and innovative** – tackles causes, not symptoms and is not afraid of experimentation
 - **Robust** – stands the test of time and works in practice from the start

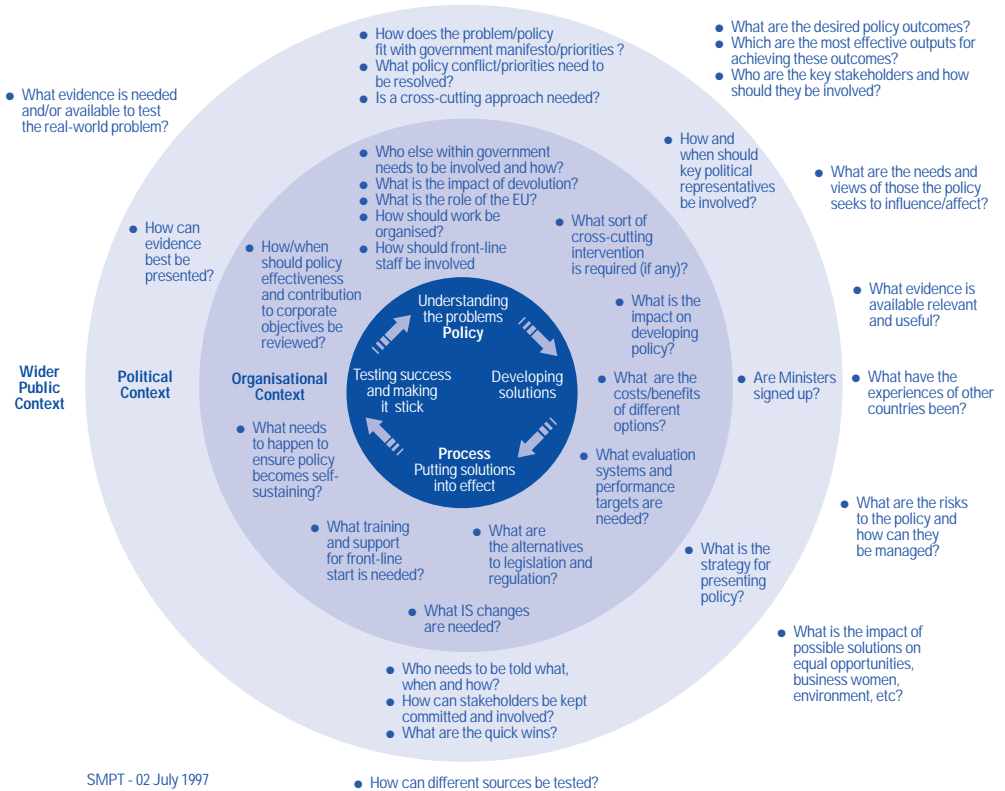
1.9 The report defines the core policy process to be:

Core Policy Process



1.10 In reality, however, policy planning is more complex and multi faceted than this and the Cabinet Office report defined “modernised” policy making in context to be more like:-

The policy process in context



1.11 Two key features of this model are – learning from successes and failures principally through more **evaluations** and developing a **policy knowledge pool** – covering, for example, the objectives of new policy projects; the results of impact assessments; relevant consultation documents and information about responses; details of evidence used; and of policy evaluation – to allow easier sharing of information about and experience of policy making and to create a more easily accessible source of evidence for future policy making.

1.12 From an independent VFM examiner's viewpoint we need to be aware of the organisational, political and wider public context of

the programmes we are examining. And we need to take these into consideration when framing the recommendations which we make arising from our VFM studies. Understanding the policy formulation process can help us to produce more insightful and genuinely constructive recommendations.

Example of how the Department of Trade and Industry (DTI) and the Department of International Development (DFID) undertake evaluations.

The DTI evaluates many of its programmes and policies yearly and publishes the results. Specialist evaluators work both throughout the department in partnership with programme budget holders and in a central evaluation unit to manage the process. Every year the DTI's central unit invites bids from the 'field' evaluators on programmes that are suitable for evaluation. The central unit co-ordinates the bidding processes and recommends the priority of each bid.

An Evaluation and Policy Improvement Committee (EPIC) has responsibility for approving the evaluation work programme for the year. The Committee is chaired by the Head of the Finance and Resource Directorate and comprises budget holders and economists.

DFID have a separate evaluation department. They select 3-4 themes from DFID's work annually for evaluation. Themes are selected for their scope to offer lessons for the future, but the aim is to cover all major sectors in a six year cycle. Within each theme, the department commissions a number of individual project evaluations. These are followed by a synthesis evaluation which combines the project-level results with other performance information, academic work and international evidence to draw general conclusions. Evaluations typically take some 15 months, using multi-disciplinary teams made up of in-house evaluators and external experts.

Who undertakes evaluations in the UK?

There are a growing number of organisations in the UK that are involved in evaluations, these are just a few:

- British Market Research Bureau International
- Centre for Applied Research in Evaluation
- Charities Evaluation Services
- Department of Environment Transport and the Regions Research Directorate
- Department of Trade and Industry Evaluation Directorate
- Economic and Social Research Council
- Environment Agency Research Directorate
- Health Education Authority
- Home Office Research, Development and Statistics Division
- Institute of Public Policy Research
- King's Fund (Health policies)
- Metropolitan Police Service (Evaluation Units)
- Market and Opinion Research International (MORI)
- National Centre for Social Research
- NOP Research Group Ltd
- Office for Standards in Education (OFSTED)
- Policy Research Unit
- Tavistock Institute
- Social Policy Research Unit

1.13 The Modernising Government White Paper (CM 4310) states that to be effective policy making must be a learning process which involves finding out from experience what works and what does not and making sure that others can learn from it too. This means that new policies must have evaluation of their effectiveness built into them from the start; established policies must be reviewed regularly to ensure that they are still delivering the desired outcome; and the lessons learned from evaluation must be available and accessible to other policy makers. The Cabinet Office Report defined the features of evaluation as:

Key characteristics:

- Evaluation should
- be systematic
 - be analytical
 - study actual effects
 - judge success

Key objectives:

- Evaluation should aim to
- improve decision making
 - help resource allocation
 - enhance accountability
 - bring organisational learning

In summary, evaluation is being given a higher profile by Government in both helping to inform policy development and in improving the knowledge pool so that good practice is much more widely identified and shared. This presents us with two challenges (i) to be able to form a judgement on the quality of evaluations undertaken by departments when appropriate and to make recommendations as to how they might be improved if necessary; and (ii) to draw on evaluations as a source of evidence.

What should you be looking for as evidence of a good evaluation?

Evaluations have a lot in common with value for money examinations and similar principles of good practice can be applied to both forms of investigation. The following bullet points bring out these common themes but also highlight the distinctive features of good evaluation:

- **Business planning.** Evaluations should be a regular feature of an organisation business planning process with “buy in” by key stakeholders.
- **Evaluability assessments** should be conducted before full evaluations are undertaken. Not all programmes are susceptible or mature enough for evaluation – management information may be too difficult to collate or the programme may be in a state of flux. The purpose of the evaluability assessment is therefore to determine what needs to be done to prepare the programme for evaluation and to diagnose the issues that warrant attention. The evaluability assessment is very similar to a preliminary study.
- **Statement of evaluation objectives.** There are many different forms of evaluation and purposes to which evaluation can be put. The underlying rationale and objectives should be clear at the outset of the examination. Objectives should also be realistic. Summative forms of evaluation are usually difficult to conduct and may need to be undertaken on a longitudinal basis, perhaps over some years. Evaluation objectives should realistically convey what the evaluation is capable of achieving.

- **Identification and inclusion of stakeholders.** Government programmes rarely involve or affect only one stakeholder group. It is important that evaluations explicitly identify who is affected by the programme from policy makers to programme managers to end-users. Steps should be taken to include stakeholders throughout the evaluation and, as far as possible, to disseminate findings and observations as they emerge.
- **Statement of programme theory.** In order to reach a judgement on the success or otherwise of a programme, evaluators need to develop a theory or model which states the conditions and actions that need to be in place for the programme to succeed. This statement is called the programme theory and might be expressed in the form of a 'Logic Model' or 'Path Analysis' setting out the steps the need to be taken.
- **Identification and piloting of methods.** The methodology used for the evaluation needs to be consistent with the overall objectives and the type of evaluation undertaken. For example, fourth generation evaluations rely more heavily on qualitative methods and these should therefore feature in the examination. All methods should be piloted ideally at the evaluability assessment stage.
- **Practical recommendations** are particularly important to evaluations since findings are intended to feed directly back into the programme. Recommendations should be clear, practical and should take full account of associated costs and benefits.

We are ready to try our fortunes To the last man

Henry IV Part Two Act IV, Scene ii



Key stages in an
evaluation and some
helpful concepts

*If you can look into the seeds of time, And say which
grain will grow and which will not, speak then to me*

Macbeth Act I, Scene iii

2.1 Evaluations follow a structured and logical approach and in this our VFM examinations are very similar. Evaluators, however, adopt different thinking and concepts to their work from which we can learn. This part discusses nine such concepts and approaches and their relevance to VFM examinations.

1 Cause and effect. Evaluations are often concerned with whether there is a clear cause and effect relationship for example, which of five health education programmes is more successful in encouraging citizens to adopt a healthier lifestyle or which employment measure in a specific geographic area has had the most success in getting people back to work. In these examples the programme which is the subject of the evaluation is the independent variable (the cause), while the planned change is the dependent variable.

2.2 Establishing cause and effect can be difficult – essentially what you are looking for are patterns in a series of events or interventions such as direct financial support or social assistance or education which can be attributed to having a direct consequence. If you can quantify or put a numerical value on the various causes and effects (outcomes) you can test the strength of their potential relationship using statistical techniques such as bivariate and multivariate analysis (The London School of Economics and Political Science's Methods Institute provide practical training for the NAO in these techniques and anyone interested in attending should telephone 020 7798 7093).

2.3 Analysis becomes more difficult when you cannot easily quantify the various factors – for example different patterns of social behaviour may be the major factor influencing whether a person responds to a health education campaign but it is clearly difficult to put a numerical value on this. In these circumstances it can help to distinguish between factors which can be clearly **attributed** to an outcome and those which only **contribute** to it (these concepts are discussed below).

- 2.4** Similarly, it may be difficult to quantify effects particularly when they are intended to be improvements in standards of living or standards of health. There are, however, techniques for awarding some form of quantified value to these sorts of outcomes such as “Quality Adjusted Life Years” .

Quality Adjusted Life Years

When examining life expectancy it may be more meaningful to weight life expectancy by some quality of life factor thus quantifying the effect of that quality factor, this is termed Quality Adjusted Life Years.

A good example is study carried out by the Stockholm School of Economics⁵ which assessed the relationship between life expectancy and income level. As life expectancy is a combination of age, gender and state of health, these factors were used to provide a Quality Adjusted Life Years value. The information required

was collected from three sources, a survey of living conditions, causes of death statistics and income tax statistics. Data were therefore available on: age; gender; state of health; life expectancy, and income levels. A regression analysis was carried out to calculate the Quality Adjusted Life Years weighting. This was done using Quality Adjusted Life Years as the dependent variable and age, gender and state of health as the independent variables. The life expectancy value was then adjusted using this weighting. The results showed a decrease in life expectancy for 20 – 29 year old men of six years in the low income category and over six and a half years in the high income category.

The difference in life expectancy for 20–29 year old males using life expectancy without adjustment and Quality Adjusted Life Years.

<i>Men aged 20–29</i>	<i>Life expectancy (years)</i>	<i>Quality Adjusted Life Years</i>
Low income	49.7	43.7
High income	53.8	47.2

Life expectancy was reduced by approximately 12 per cent when Quality Adjusted Life Years weighting are applied to the data.

⁵ Income-Related Inequality in Life-Years and Quality-Adjusted Life-Years. Ulf-G Gerdtham* and Magnus Johannesson, Stockholm School of Economics, September 1999.

2.5 How is this relevant to VFM examinations? In designing studies we should start by ensuring that we have a clear understanding of the output or outcome – “ the effect” which the programme or project we are examining is intended to achieve. Having done so we need to “unpack” the programme and identify the factors “ the causes” which will ensure the success or failure of the programme. Value for Money will often depend on how well these factors are managed and how unforeseen or risk factors are dealt with. Key issues we might focus on are:

- Could the causal factors be better managed and delivered in a way which would improve the quality of the output or outcome; increase its quantity and its accessibility to those intended to benefit; or improve the speed of delivery?
- Are the risks likely to lead to failure or sub optimisation of programme delivery identified and well managed?

2 **“Attribution” and “contribution”**. In any cause and effect relationship there will be factors which can be clearly attributed to the intended outcome or output. Such factors are identified by asking the question if they did not exist or were not provided would the output or outcome be delivered. If the answer is no they are probably attribution factors. Other factors might influence the success or failure of the outcome but the strength of the relationship might be weaker and therefore they may be more contributory factors. For example, take the hypothesis that lack of employment among released prisoners results in a return to crime (recidivism). The key attributable factor is likely to be lack of a job but contributory factors may be lack of appropriate social support or counselling to readjust to life outside prison. Evaluators often make the distinction between attribution and contribution in analysing success or failure and in the case of the latter the action needed to remedy it.

2.6 How is this relevant to VFM examinations? It is very relevant for the following reasons:

- It can help in examining joined up initiatives by identifying which departments and organisations are likely to have more influence on the success of a joined up initiative.
- Distinguishing between attribution and contribution can help us target our evidence collection. Clearly if we are going to say that something is clearly attributable to the success or failure of a programme the evidence will need to be much stronger than if it is only a contributory factor.
- It can help ensure that our recommendations focus on improvements (more likely to be “attributable” interventions) which are going to have most impact.
- It may improve the potential for financial impact – improving the management of factors directly attributable to an output is more likely to deliver quantified improvements than focusing on contributory factors (although these may be important).
- It can help demonstrate to those who read our reports that we have a good understanding of the complex programmes which we often examine. Clearly we cannot ignore contributory factors but distinguishing between attribution and contribution can help improve our understanding and analysis.

3 Modelling. Evaluators often develop models to help them identify the complex interrelationships that exist between inputs, outputs and outcomes. Models can range from the simple to the very complex. Their main advantage is that they allow you to consider in a systematic and objective way all the factors that will influence a particular outcome. They can help explain why things happen in a certain way, predict future events or outcomes under a range of different assumptions, and help in diagnosing why a desired outcome was not achieved. Models have some limitations – they are only as good as the data on which they are based; they should not be impenetrable black boxes; and modelling behavioural aspects is often difficult.

Some models used by Government

The Home Office property crime models attempt to explain the level of property crime in terms of demographic and macro economic factors.

The DETR National Road Traffic forecasting Model is a statistical representation of the capacity of the road system.

OFWAT's efficiency models are used to model outputs in setting efficiency targets for water companies.

Lord Chancellors Departments modelling of civil legal aid expenditure gives an indicative rather than precise estimate of future spending on legal aid.

Source: Adding it up Improving Analysis and Modelling in Central Government Cabinet Office January 2000.

How is Modelling relevant to VFM examinations? Modelling is useful in our VFM work for two main reasons:

- i **As a diagnostic tool.** We can use simple models (logic models are explained in part 3) to identify in a structured way all the key input variables which form part of a programme or project. Understanding how a programme or project works in detail should be a key part of the study design process so that you can fine tune and focus your analysis where it will lead to most added value.

- ii **In demonstrating financial impact.** By developing simple models we can show in our reports how changing certain assumptions might improve efficiency or generate savings. A good example of this was our report *Improving Energy Efficiency Financed by a Charge on Customers* (HC 1006 1997 – 98).

Improving Energy Efficiency Financed by a Charge on Customers (HC 1006 1997 – 98)

We modelled the impact of changing the mix of energy efficiency projects financed by the scheme and also the impact of improving the cost effectiveness of projects. The model indicated that the scheme could achieve more for customers either by increasing energy savings by up to around 9 per cent equivalent to insulating the homes of an extra 8000 customers a year or achieve a larger reduction in customers' electricity bills by up to £40 million a year. In addition to these two options, benchmarking indicated that improving the performance of projects to match the cost-effectiveness of the top 50 per cent would increase the financial benefit to customers by some £25 million between 1998 and 2000.

4 Different paradigms. A paradigm is a model or set of assumptions which reflect a particular way of conceptualising given situations. Thus a social scientist may adopt a different paradigm in attempting to explain cause and effect compared to the paradigm adopted by an economist or auditor. This means that the paradigm which you view a situation from will influence the analytical methodologies which you will probably select. There is always the risk, however, that the paradigm which influences your thinking can suppress new lines of inquiry or stifle creativity. A typical paradigm debate is the balance in the influence accorded to behavioural human responses and direct programme interventions on the human impact of the government's social policies.

How is this relevant to VFM examinations? Firstly, it is useful to know the paradigm underpinning a programme – is it driven predominately by economic or social thinking. This can help understand the context of activity and ensure that our recommendations reflect this context if appropriate. Secondly in designing studies it can help in looking at the key issues from the perspective of different paradigms and to synthesise these so that the focus of your examination is sufficiently comprehensive. A good example is the different approaches which an economist and an auditor are likely to apply to definitions of efficiency. An economist is more likely to consider wider issues such as the benefit foregone from a certain action – the opportunity cost and the displacement effect.

5 Behavioural aspects and organisational effectiveness.

Evaluations often focus on the behaviour of individuals as a key factor influencing the output and outcome of a programme. Evaluations may also explore the culture of organisations to assess its influence on outcomes. Organisational effectiveness and how to achieve behavioural change to promote, for example, less risk adverse management and more innovation are key issues being addressed by the Modernising Government programme.

Evaluating behaviour and organisational effectiveness is however not easy and the evidence available is usually regarded as much “softer”. It is possible, however, to form a structured assessment of such behavioural aspects as an organisation's capacity to change and respond to innovation. A good example is the work of the Institute of Local Government Studies of the University of Birmingham (inlogov).

- 2.7** Inlogov defined four key characteristics where were likely to demonstrate an organisations' capacity to change (set out below).

Characteristics likely to suggest an organisation's capacity to change and adapt

1 Capacity to adapt to external forces

- Responsiveness to users, citizens and communities
- Effectiveness of networks and partnerships as triggers for change

2 Capacity to deliver business results

- extent of innovation in business processes
- effectiveness of management systems (strategy, service planning, delegation etc)

3 Capacity for accountability and control

- extent of focus on probity and accountability
- extent of focus on performance evaluation

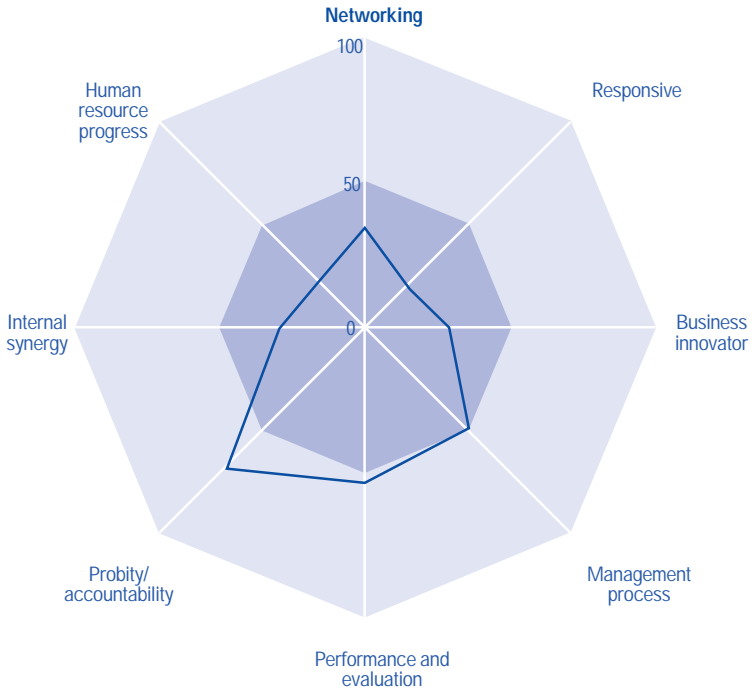
4 Developing cultural capacity for the future

- effectiveness of human resource management
- extent of focus on developing internal synergy (eg cross department working)

- 2.8** These characteristics can be used as a diagnostic tool in the form of a more detail questionnaire by which managers rate their own organisation's – performance under the four headings. The responses to the questionnaire can then be plotted to form an overall picture of an organisation's culture. Two examples of different results are shown overleaf.

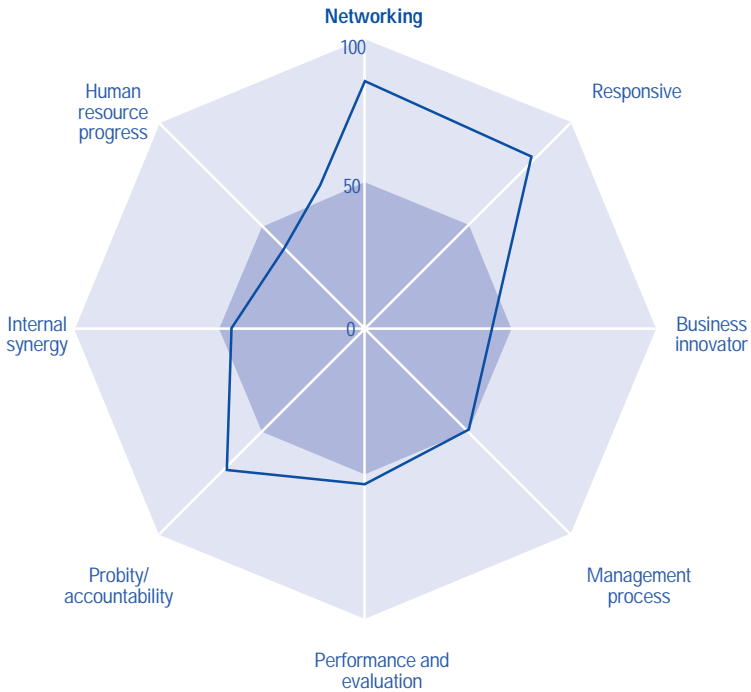
- 2.9 How relevant is this to VFM examinations?** This sort of analysis is based on independent responses to a questionnaire and is a legitimate source of evidence as long as the purpose of the responses being given is fully understood by those being consulted and the analysis of the data can be undertaken in a structured way.

Organisation A



The graph suggests a traditional bureaucracy. The focus is on strong adherence to internal rules and norms – on how things are to be done, rather than what is to be delivered. While the focus on probity and accountability can be viewed as a strength, this is not well balanced at present against external networking and responsiveness. The low scores on internal synergy suggest the legacies of strong departmentalism, and staff perceive the organisation to be weak on human resource issues. Relationships with communities, partners and other external sources of innovation are viewed as weak, though networking is beginning to develop.

Organisation B



This organisation has a strong tradition of innovation over an extensive period of time. The high external orientation scores reflect a long period of partnership working and community consultation and engagement, and the organisation has also been proactive in developing innovatory approaches to service design and delivery and management systems and processes.

Note: The further the dark line is away from the centre the more the organisation meets the characteristics.

Source: Institute of Local Government Studies (INLOGOV), The University of Birmingham.

2.10 As source of evidence this approach may be more useful in diagnosing organisations' performance rather than making specific recommendations. And its conclusions would have to be substantiated by other evidence such as business results and achieving key targets and outputs. But it is, nevertheless, a useful concept for identifying issues worth exploring in more depth to form an objective assessment of organisational effectiveness or to explain under performance.

6 **Longitudinal analysis** traces the life histories of projects, programmes and individuals across time and this may be for many years. Because data is collected and analysed over a long period longitudinal analysis is good at establishing causal links between inputs, outputs and outcomes which a point in time analysis may not identify. Longitudinal analysis can also help to reduce the risk of "compensation effects". For example there may be extraneous variables which result in outcome effects that obscure or exaggerate the "true" effects of inputs. Similarly there may be "stochastic" effects – measurement fluctuations attributable to chance. Longitudinal analysis can help to minimise these factors and the risk of a false conclusion because data is collected over a much longer time period.

A good example of longitudinal analysis would be where the government had decided to implement new measures to reduce juvenile crime and data on their impact, for example the number and type of incidents, the number of arrests and re-offenders, were collected over a five year period to compare trends before and after the new measures were in place.

2.11 How relevant is this to VFM examinations? Reliable data analysis is critical for our VFM work but commissioning longitudinal data can be expensive and difficult to acquire within the time horizon of a typical NAO study. When we commission independent data collections through surveys for example, this is covering a much

shorter time period. The principle of analysing data over a longer time period is, however, very relevant to our work and we should seek opportunities to do this particularly when such data is available in departments. Trend and time series analysis as well as more complex analysis of the relationship between inputs, outputs and outcomes all become possible providing more depth to our reports and their recommendations. In tracking progress over time our reports on privatisation, information technology projects and major defence equipment projects are a form of longitudinal analysis.

7 Experimental design. In seeking to form a judgement on cause and effect an evaluator will often adopt an experimental design approach. This involves establishing “experimental” groups to whom some form of intervention is delivered for example, education to help unemployed youths get jobs. The experimental group is compared to a control group who are not exposed to the intervention or to a comparison group who experience a different intervention. A simple example of this type of approach might be gauging the effectiveness of educational training films by showing alternative versions of training films to randomly selected employees whose understanding of the lessons learned is then measured before and after viewing. The control group would be the group of employees who see no training film.

Experimental design can be very sophisticated – with those in the experimental and control groups being selected randomly so that any observed difference in outcomes can be attributed to the effects of the intervention and not to any differences in the composition of the groups. There is a risk, however, in that experimental designs are often used to evaluate the impact of different social programmes and their success will often depend on the behavioural response of those the programme is intended to benefit. An experimental design will not necessarily allow you to take account of behavioural aspects which may be critical to achieving desired outcomes.

2.12 How relevant is this to VFM examinations? Experimental design is not likely to be a technique which we would use directly in a VFM study. We might, however, use a form of it particularly in examining a series of case studies of different approaches adopted by departments to address a similar problem.

2.13 We need to be aware of experimental design because as departments tackle longer term intractable social problems – “Wicked issues” (explained in the introduction to this guide) they are experimenting with policy interventions and measures often in pilot projects with a requirement to evaluate their effectiveness before deciding whether to implement them more widely. Depending on the time and resources available experimental design may play a wider role in these evaluations and we need to know of both its strengths and weaknesses as an evaluative tool.

8 Sustainability Evaluations often focus on the sustainability of outcomes achieved by programmes. This will involve assessing whether all the factors – training, infrastructure, appropriate incentives, changes in behaviour by those intended to benefit from the programme are all in place or adequately addressed to ensure that desired outcomes are likely to be sustained over the desired time horizon and do not become a one off event.

2.14 How relevant is this to VFM examinations? Increasingly as our VFM studies examine output and outcome achievement we will need to form judgements of their likely sustainability in the medium to longer term. To do so we will need to know the critical factors which are likely to promote sustainability. Sustainability is an aspect we should take into account in framing the recommendations which arise from our studies.

9 Deduction or Induction. Evaluations can take different approaches to assess programme effectiveness. With a **deductive** approach the evaluator will decide in advance what will constitute a successful programme outcome. The aim will then be to measure the attainment of these outcomes using a methodology which will provide as precise as is practicable attribution of inputs to outcomes. With an **inductive** approach the evaluator will collect data on the programme's outcomes and impacts and build-up an 'evidence-based' understanding of how outcomes can be related to inputs.

2.15 How relevant is this to VFM examinations? The traditional approach which we adopt for a VFM study is to start with the objectives of a programme or project. While this is logical it can result in a narrow focus for example, assessing the processes needed to achieve the objective rather than those intended to benefit from the programme. An inductive approach may be better in some circumstances, particularly when we want the study to assess effectiveness from the perspective of those intended to benefit from it. This is very much in keeping with the commitment of the Modernising Government programme that public services should be designed very much with the end recipient in mind.


2.16 To recap:

Concept	When is it likely to be most useful
1 Cause and effect	Examination of outputs and outcomes, response to “Wicked” issues; evidence of poor programme delivery or sub optimisation requiring investigation.
2 Attribution and contribution	Testing the strength of evidence and its relationship with outputs and outcomes. The concept is also useful for considering financial benefit potential.
3 Modelling	Seeking to explain complex programmes and processes and the relationship between cause and effect.
4 Different paradigms	Testing the strength of your study focus and considering the scope for different perspectives which might lead to new lines of analysis which add value.
5 Behavioural aspects	Key to assessing organisational effectiveness, quality of service examinations and evaluating change management.
6 Longitudinal analysis	Seeking to explain trends in the delivery of outputs and attainment of outcomes.
7 Experimental design	Important for establishing causal links in the attainment of outcomes but more likely to be a source of evidence than an approach we would use directly as part of a VFM study.
8 Sustainability	Longer term consideration of the impact of government programmes.
9 Deduction or Induction	Important for designing studies and considering new ideas and approaches.

In summary, evaluations tend to adopt a wider range of concepts and perspectives than value for money examinations in seeking to form objective assessments of why things have happened in a certain way. Evaluations also give more emphasis to social behavioural aspects although the evidence to form judgements on these is inevitably softer. Adopting some of the concepts which evaluators use can help as to think more laterally in VFM studies to help us to get to the heart of an issue and make recommendations which add real value.

*The eye sees not itself
But by reflection, by some other things*

Julius Caesar Act I, Scene ii

A man in a dark suit is shown from the chest up, holding a human skull in his right hand. He is looking down at the skull with a somber expression. The entire image is overlaid with a semi-transparent red color. The background is dark and indistinct.

Evaluative techniques
and how they can help
in VFM examinations

*We know what we are,
but we know not what we may be*

Hamlet Act IV, Scene v

- 3.1** This part explains ten techniques which are often used by evaluators and which you may find useful in carrying out VFM studies. Remember, however, that it is important from the outset to distinguish between diagnosis and analysis:

Diagnosis. This is the initial stage when you assess the potential for improving VFM or ascertaining whether under performance exists. At this point the techniques which you use should help you to diagnose the potential issues to home in on or the areas where you should concentrate most of your analysis. Diagnosis may require techniques that are less useful for more detailed analysis.

Analysis. This is the main focus of a VFM examination and is concerned with understanding why things happen as they do and the scope for improvement.

Ten Techniques

1. Meta Evaluation and Meta Analysis

- 3.2** Many of the programmes or projects which we examine may have been subject to many separate internal and external reviews – these may be formal evaluations by academics, think tanks, lobby groups or stakeholders or efficiency or inspection reviews. If our work is to be comprehensive and authoritative it is important that we draw on these reviews when they are of high quality and can contribute to our analysis – it can also be cost effective to do so. We need to know whether any particular evaluation or reviews influenced the design or impact of the programme we are examining and form a view as to whether we can rely on the evaluation as a source of evidence. Access to these sorts of evaluations and reviews is made much easier by the Internet and Government Secure Intranet (GSI). We may also decide that we can rely on the data in a series of evaluations and in which case we will have analyse this in a systematic way – Meta Evaluation and Meta analysis can help do this.

- **Meta Evaluation** is a technique for summarising and analysing the results of a series of evaluations. It can be particularly useful because the findings of a single evaluation in isolation may miss fundamental issues or be contradictory. Meta Evaluation allows you to assess a number of evaluations as a single body of evidence thus reducing the risk of fundamental issues being missed and allowing you to reach a more authoritative assessment of the validity or strength of contradictory messages coming from different single evaluations. There are a number of criteria you need to apply in assessing a series of evaluations – these include reliability of the evidence on which the evaluations are based; when the evaluations were completed – are the results still valid; the quality and expertise of those who undertook the evaluation; the level of acceptance of the evaluations’ recommendations; and the action taken as a result. This analysis is best presented in a matrix format.

- **Meta Analysis** is a more sophisticated technique. Evaluations may be based on a range of quantified data and it is sometimes possible to analyse this data systematically to arrive at valid conclusions. For example, there may be a series of evaluations over a number of years on the impact of training measures to reduce adult illiteracy. It can be valid to treat each evaluation as a case and attempt a multi variate analysis (this technique is explained below) of the different inputs evaluated to assess their impact on adult literacy to determine the net impact of the various training measures. A more complex example is the International Cochrane Collaboration which was set up in 1992 to promote the collation and dissemination of evidence from randomised controlled trials of new medicines and drugs, and aims to establish a register of all trials with a view of producing regular meta analysis.

- 3.3** A technical note is available on MERLIN which explains in a practical way how to use Meta Evaluation and Meta Analysis and where to go for expert advice and help.

2. Organisational Mapping

- 3.4** Many organisations may be involved in delivering a government programme and even within one department there may be many divisions or agencies who can influence the success of a programme. Good examples are the criminal justice system where the Crown Prosecution Service, the Courts, the Police, Legal Aid, Probation service, the Prison Service all have complementary roles, and the Government's initiatives to improve the performance the construction industry which involves many private and public agencies. Value for money will depend on how well those departments and agencies involved in a programme are organised to avoid duplication, to ensure that services are complimentary, to provide a consistent and easily understood interface with clients, and above all that they are working in unison to achieve a common outcome that is sustainable. It is also important that organisations have in place the right incentives and business processes to promote effective partnering and joint working.
- 3.5** In assessing VFM we often have to form a judgement as to how well those involved in programme implementation are organised and organisational mapping is a useful tool for doing this. It involves:

Step one – identifying all organisations or parts of an organisation which can influence the success or failure of the programme or project you are examining.

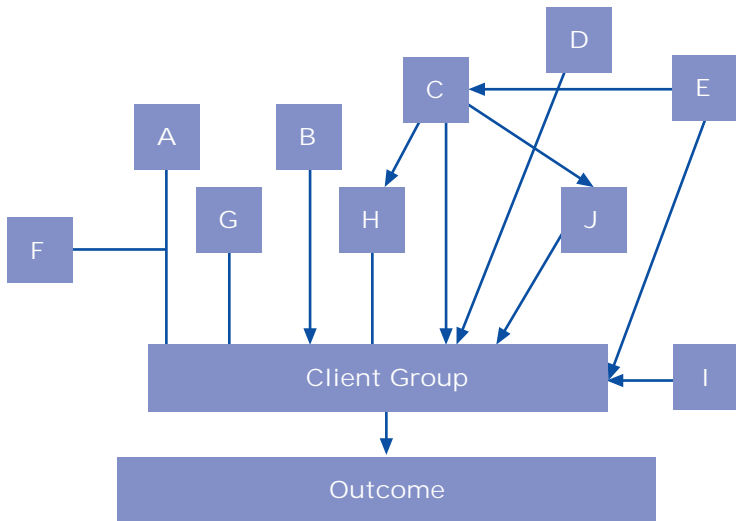
Step two – taking each key objective of the programme or project and assessing the contribution which each agency or department is likely to make to this. It may help to rank the contribution as high, medium or low.

Step three – producing an “organisational map” showing diagrammatically the linkages between each organisation and the key outcomes.

Step four – assess using the “organisational map” whether each organisations’ involvement appears relevant, consistent and complimentary.

Step five – investigate apparent overlaps; contradictory impacts for example, where one organisation’s input appears inconsistent with another’s; where those most effected by the desired outcome are unlikely to seek advice because of an unclear client/customer interface; or where overall programme delivery appears incoherent and badly organised putting outcome delivery at risk.

Example of an organisational map suggesting duplication of organisational roles in a programme and unclear client/customer interface



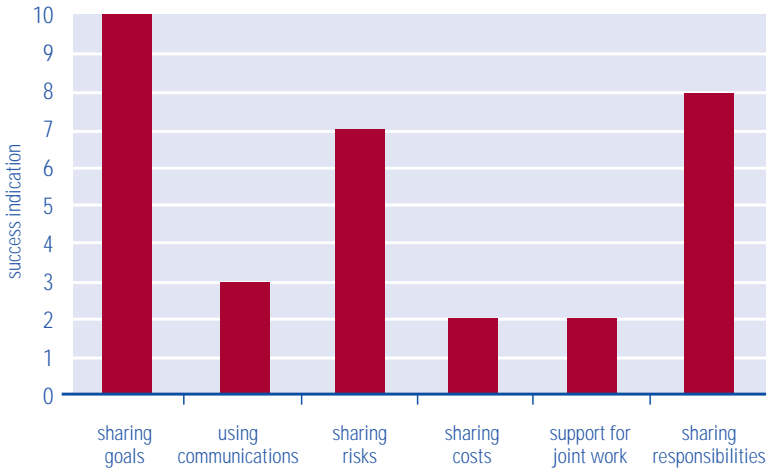
Organisational Mapping can also be used in a directed sense to gauge the strength of partnering arrangements (or 'Joined-Upness') between two or more organisations. The approach resembles that set out in Steps 1 – 5, in that organisational relationships are mapped-out and assessments are undertaken of the strength and impact of those relationships on outcomes. The quality of partnering arrangements however can be examined in relation to pre-determined criteria for effective joint-working (either set by the evaluator

or the organisations examined). Each of these criteria are then explored by using a mixture of surveys, focus groups and Cognitive Mapping (explained later) in order to determine a score for each criterion. For example, where it is agreed that the best conditions for joint-working are where goals of different organisations completely match one another (carrying a score of 10), and the worst conditions are where goals conflict (carrying a score of 0), we can draw on focus groups and surveys to pinpoint where the organisations examined feature on each scale. This helps to systematically highlight strengths and weaknesses. An example of how this might be presented is shown overleaf.

Example of 'partnering' criteria used with Organisational Mapping

- sharing of goals
- clarity of communication
- risk sharing
- degree of joint funding
- support for joint working
- shared responsibility and accountability

The extent to which joint working is successful



3. Web-based data collection tools⁶.

3.7 The growth in the Web – the Internet is a major new medium for collecting data quickly and cost effectively. We can use the Internet in various ways to collect evidence:

- **E-mail questionnaire.** The questionnaire is sent to a list of known e-mail addresses who respond electronically. You do not need to be on-line or directly connected to the respondent at the time the reply is sent. Software can be used to prepare the questionnaire, the e-mail address list, and to extract the data from the replies. The advantage is that by showing up in a respondent's e-mail box, e-mail questionnaire's demand immediate attention. The disadvantage is that they are likely to have to be short and limited text.

⁶A good source of information is an article by James H Watt – Internet Systems for Evaluation Research in New Directions for Evaluation number 84 Winter 1999 – A publication of the American Evaluation Association.

- **Web survey Systems** are software systems specifically designed for Web questionnaire construction and delivery. They combine the administration tools of computer aided telephone interviewing (CATI) with the language of the Internet. Web survey systems allow you to create complex questionnaires that are visually appealing.
- **One-on-one text based Internet interviews.** In its simplest form two people connect their computers by knowing each other's unique ID. The interviewer and respondent interact by typing questions and answers in a text area that is viewed by both parties. The interview is conducted in a standard back and forth manner. This approach can be expanded to include "chat rooms" for interaction by which a number of participants can respond to the interviews; and to include an audio visual facility requiring a video phone that includes a small camera and micro phone.
- **Web based focus groups and text-based focus groups.** These duplicate some of the moderated group interaction of a traditional face-to-face focus group discussion. Web focus groups can be conducted with text interaction (the most common method) or by using audio-visual capabilities of the Internet.

3.8 The Internet can also be used to test the quality and responsiveness of organisations' web sites. For example, the Report 'Government on the Web' (HC 87, 1999 – 2000) used a non-reactive census of all government Web sites (www.governmentontheweb.org) to assess the key features of each site. Using 'mystery shopping' (paragraph 3.30) you can assess the quality of public services and information provided through web-based to the citizen.

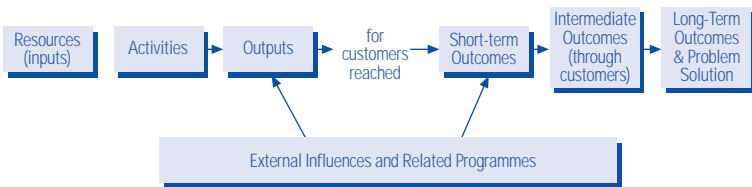
3.9 For further advice on Web-based data collection tools contact Alex Scharaschkin and Alison Langham on 020 7798 7171.

4. Logic Models

3.10 Evaluations and VFM studies are about understanding how a programme or project works and having collected sufficient information on programme implementation determining whether the programme or project can be improved or underperformance remedied. To do all this we need tools that help us diagnose the factors and various stages involved in a programme and logic models are one such tool.

3.11 A logic model breaks a programme down into its constituent parts. In its simplest form a logic model is

Elements of the Logic Model



Source J A McLaughlin and G B Jordan. Evaluation and Program Planning February 1999
ISSN 0149-7189

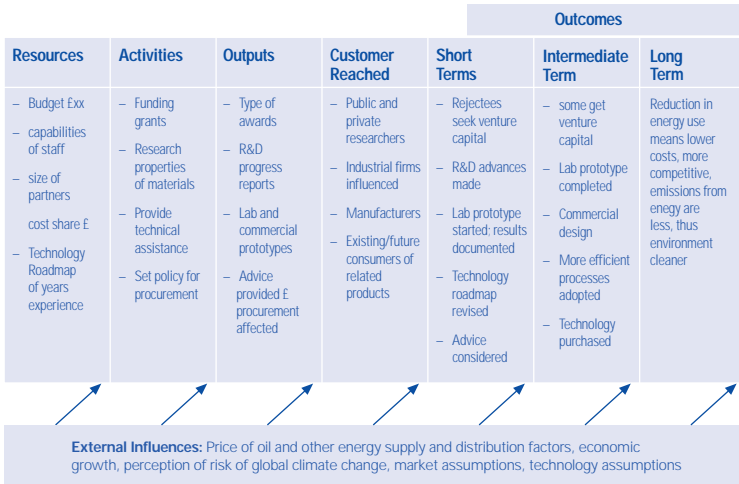
3.12 The benefits of using a logic model are:

- Allows you to build a good understanding of programme and expectations for resources, customers reached and results. Agreeing the logic model with the client department can be useful because it ensures that you have common understanding of how the programme is intended to work and what the key inputs, outputs and outcomes are supposed to be.
- Helps you identify projects or programmes that need to be improved or components (for example human and financial resources, partnerships, grants) which are critical to achieving outputs and outcomes, or where the various inputs and outputs have inconsistent or implausible linkages.

- Assists you in identifying key performance measurement points and evaluation issues making it easier to know where to target your data collection.

3.13 A critical feature of a logic model is the identification and description of key contextual factors external to the programme and not under its control that could influence its success either positively or negatively. It is important to examine the external conditions under which a programme is implemented and how those conditions affect outcomes. This explanation helps clarify the assumptions on which performance expectations are set.

3.14 A more complex example of a logic model is shown below based on an energy technology programme⁷.



A table with elements of the Logic Model for an energy technology programme.

⁷J. A. McLaughlin and G. B. Jordan. Evaluation and Program Planning, February 1999. ISSN 0149-7189.

5. Cost benefit/effectiveness Analysis

3.15 Cost benefit analysis involves estimating the benefits of a programme both tangible (for example, number of patients treated) and intangible (for example, changes in social behaviour), and the costs of undertaking the programme – both direct and indirect. Once specified, the benefits and costs are translated into a common measure, usually a monetary unit. Clearly the assumptions underlying the definitions of, and the measure of, costs and benefits strongly influence the resulting conclusions consequently you need at the very least to state the assumptions underlying your analysis.

3.16 Cost benefit analysis is easier to use when applied to industrial and technical projects, where it is relatively easy to place a monetary value on benefits as well as costs – for example, engineering projects intended to reduce the costs of electricity to consumers, road construction to improve transport or irrigation to improve agricultural yields. Estimating benefits in monetary terms is more difficult with social welfare programmes where only some of the inputs or outputs may lend themselves to monetary quantification – for example, it may be possible to quantify the increase in those securing employment from attending specific training but there may well be other longer term gains such as improving social cohesion and quality of life which are impossible to quantify precisely.

3.17 Cost effectiveness analysis involves putting a monetary value on programme costs only; benefits are expressed in some form of other output unit. For example, the cost effectiveness of distributing free textbooks to rural primary schoolchildren could be expressed as follows – each 1000 project £s increased reading levels by an average of one grade level⁸.

⁸Evaluation A Systematic Approach Peter H. Rossi and Howard E. Freeman.

- 3.18** Cost benefit analysis and cost effectiveness analysis can be quite technical to use and you will need to seek expert help. There are various circumstances when this sort of analysis may be important in a VFM examination – (i) You may be examining the implementation of a programme or project and you may need to have assurance that the type of programme intervention adopted was subject to proper critical appraisal before being selected; in this case you will want to know that the cost benefit or cost effectiveness analysis was applied to a good standard (ii) In forming a view of the success of a programme in achieving its planned outputs or outcomes you may decide to carry out your own cost benefit or effectiveness analysis.
- 3.19** It is important never to forget that any programme intervention or project may have external or “slipover” effects, that is, side effects or unintended consequences that may be either beneficial or detrimental. Given that such effects are not deliberate outcomes, they may be omitted inappropriately from cost benefit calculations unless special efforts are not made to include them. The most common negative external effects of industrial projects are pollution, noise, traffic, and destruction of plant and animal life. Externalities may be difficult to identify and measure but, once found should be incorporated in to the cost benefit calculations.

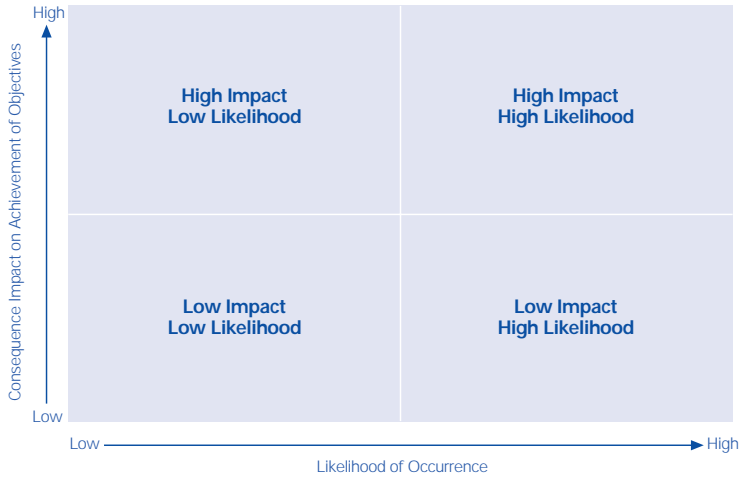
6. Risk Assessment

- 3.20** Much of our work is about assessing the risks to value for money. We often need to examine the quality of organisation’s business risk management process to identify areas for examination or we evaluate organisation’s own risk assessments as part of a wider review of a project or programme. Risk assessment can assist in identifying whether the resources devoted to managing risk are consistent with the level of risk to the organisation’s objectives in terms of the responsiveness and quality of public services and the delivery of desired outcomes.

3.21 Risk assessment is particularly useful as a diagnostic tool at the problem formulation stage. It consists of a number of stages:

- **Step 1: Clarity of objectives.** Establishing whether departments have clear statements, aims, objectives and plans for delivery of policy outputs, services and outcomes, and ensuring that these are communicated throughout the organisation.
- **Step 2: Identification of risk.** Examining whether departments identify the key risks for which they are responsible and those risks which are most likely to impact on performance and delivery of public services.
- **Step 3: Assessment of risk.** This will involve *analysis* and *evaluation* of risks to provide an overall assessment of the potential impact of identified risks, and the timescale over which the risks need to be managed.
- *Analysis* should determine existing controls, and establish consequence and likelihood of risk in the context of those controls. Consequence and likelihood may be combined to produce estimated level of risks, quantified wherever possible, or qualified in terms of a range from low to high.
- *Evaluation* of analysed risks then enables risks to be ranked so as to identify management priorities and present information for business decisions about which risks need to be addressed (for examples, with a major potential impact and a high likelihood of maturing).

Example of Analysis of Risk



- **Step 4: Treatment of risk.** Determining the resources departments devote to manage risk, and prioritising and allocating responsibility for risk. For example, low priority risks may be accepted and monitored to keep an eye on them. Treatment of higher priority risks may require adopting alternative means of project, programme, or service delivery.
- **Step 5: Monitoring and Review.** Whether the department have a continuous process which must include monitoring and reviewing the identified risks, and being open to new or changed risks and opportunities resulting from evolving circumstances.

Assessment of risk does not of course mean that events over which an organisation has little or no control will not come out of the blue but VFM examinations should be able to draw conclusions about whether management have adopted a well thought through approach to risk management, risk taking and innovation.

7. Cognitive Mapping

3.22 Cognitive Mapping is a versatile interviewing technique suited to both problem formulation (or diagnosis) and analysis. It was first developed as a soft Operational Research technique as a means of systematically understanding how key individuals perceive the operation of a particular programme, project or activity. For evaluation it is a very useful tool for all forms of analysis, in particular Fourth Generation Evaluation which aims to understand a programme through the eyes of the key stakeholders (discussed in Part 1). Cognitive Mapping has a number of strengths which make it an invaluable tool in VFM examinations and evaluations:

- It enables the views expressed by interviewees to be visualised and effectively 'cleared' at the point of data capture
- It is an efficient way of focusing on the strengths and weaknesses of the programme(s) examined
- It encourages problems to be seen from the perspective of those involved with and affected by the programme(s) examined. It can also be used specifically to target perceptions of contributions to programme outcomes.

3.23 Cognitive Mapping can be broken down into four distinct stages:

- **Step One.** Identify the key players who have the greatest influence over, or are affected to the largest extent by, the programme examined.
- **Step Two.** Either individually or in focus groups, undertake brainstorming sessions with the key players in order to build-up a picture of how they perceive the programme examined and the incentives, concerns, strengths and difficulties relating to the programme evaluated. If the evaluation is outcome-focused you will need to develop an understanding of how these factors contribute to an outcome.

- **Step Three.** With the key players interviewed begin to order and graphically represent the views expressed. This should be either on a large whiteboard/flipchart or increasingly with the assistance of bespoke software such as 'DecisionExplore'. The result should be a logical map setting out the factors that exercise an influence over the running of the programme. Where concerns or difficulties are identified you may wish to ask interviewees under which conditions perceived difficulties might be solved, or considered acceptable.
- **Step Four.** Analyse the complete map(s) with a view to identifying recurrent strengths and weaknesses, overlaps and gaps between different players within the programme.

3.24 Guidance is available on Cognitive Mapping in the Qualitative Handbook 'Getting Beneath the Surface'. A demonstration of the software package DecisionExplore can be downloaded from the website <http://www.soc.surrey.ac.uk/caqdas/> and external advice can be provided by Professor Valerie Belton at the University of Strathclyde.

8. Multi-Criteria (Decision) Analysis

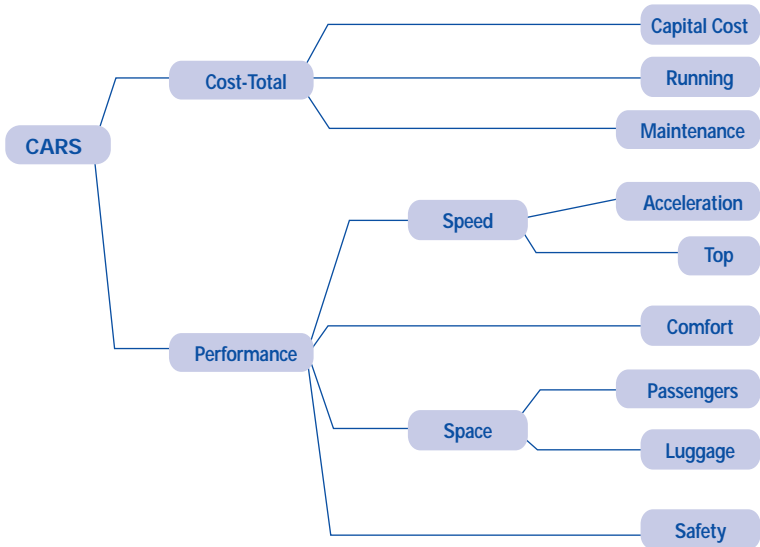
3.25 Many of the programmes that we examine in our VFM work begin within decisions taken by Departments and Agencies on the best way to achieve certain goals or outcomes. Typically we have looked at topics as wide ranging as procurement decisions, the prioritisation of project funding and the logistics of military spares. Multi-Criteria (Decision) Analysis is a formal technique designed specifically to explore how decisions are made and to reflect on whether the right apparatus is in place to make the best decisions.

3.26 Multi-Criteria Analysis is a model building approach to analysis that requires key decision-makers to come together and through a process of facilitation to construct a framework for decision-making. These days there are computer packages, for example VISA, that facilitate

the model-building process and enable adjustments and sensitivity analysis to be conducted with great ease. Nevertheless this can still be undertaken without tailored software packages.

3.27 In its simplest form, the technique involves identifying the criteria of a decision and the relative importance of each criterion. For example, if a procurement decision was being made on the purchase of two competing X-Ray machines, the assembled key players might identify, for the sake of simplicity, size, reliability and speed of photography as three criteria, carrying weights of 20%, 30% and 50% respectively. Competing options would then be scored against the model under each of the criteria identified and multiplied by their respective weights. The result would then be a single, robust score for each of the options under consideration. An illustration, using the example of a decision to buy a car, is shown below.

An example of multi-criteria analysis as applied to a decision to



- 3.28** This approach can be used in two main ways: to assist decision-making or as a basis for assessing the strength of past decisions. In both cases the result can be a firmer and more visible framework for decision-making and provides an evaluation benchmark to return to in later years.
- 3.29** A Technical Note is available on MERLIN explaining the practicalities of using Multi-Criteria Analysis and the sources of advice available.

9. Mystery Shopping

- 3.30** Within the Market and Social Research community Mystery Shopping has been a long established technique. It is based on the principle that the best way to judge the quality of a service or to understand its effect on users, is to gain first hand experience as a consumer of the service examined. What distinguishes Mystery Shopping from any other form of participant observation, is the evaluator's decision not to reveal themselves as an evaluator. In every other respect therefore the Mystery Shopper acts and behaves like a genuine service recipient.
- 3.31** In the context of evaluation, Mystery Shopping provides one means of accessing the outcomes of a particular service or process. It also fits well with particular forms of evaluation, such as the goal free or 'black box' evaluation, which de-emphasises the importance of programme or service objectives, in favour of the direct impact on those affected by the activity examined. In this sense therefore it is a technique that marks a departure from the 'normative' tradition of value for money work.
- 3.32** The technique may need to be handled with care as it requires a degree of duplicity on the part of the evaluator and the organisation they belong to. With VFM examinations, arrangements may need to be put in place to inform, perhaps at a high level, the intention to Mystery Shop. Extra care may also need to be taken at the sample selection stage to ensure that sites visited (whether real or virtual, as in the Government on the Web HC 87 1999 – 2000 study) are selected

to be representative. In most cases however the evidence provided by Mystery Shopping can be a valuable measure of the quality of service provided and can provide a unique insight into the outcomes of programmes and processes.

- 3.33** For further information on Mystery Shopping and good practice in its application you may wish to contact the Market Research Society who have developed a very useful code of conduct on the subject. Their website address is **www.marketresearch.org.uk**

10. Multi-Variate Analysis

- 3.34** Multi-Variate Analysis is a statistical technique which is particularly well suited to assessing the relationship between inputs and outputs or outcomes. For this reason it is a key technique in the evaluators toolkit and, if used advisedly, can be a strong source of evidence in any outcome-focused evaluation.
- 3.35** Multi-Variate Analysis is the term given to a package of statistical techniques that help to chart the relationship between many variables and a single output or outcome variable. An example might be an examination that sought to explain the level of crime in different areas of the country. The crime level would be defined as the outcome (or response) variable and possible inputs (defined as explanatory variables) might include expenditure on policing, number of police on the beat, number of police vehicles, use of firearms, use of community policing etc. Under these circumstances Multiple Regression – which is a branch of Multi-Variate Analysis – might be used to determine which of the explanatory variables selected have an influence on programme outcomes. It might be found, for example, that a programmes to reduce crime through the introduction of community policing may be far more effective than increased numbers of police on the beat. This would help to inform Police divisions and related Departments and Agencies, from a Summative perspective, where efforts would be best placed to deliver benefits.

- 3.36** A Technical Note is available on Multi-Variate Analysis along with advice from the Technical Advisory Group (contact Robin Ryde x7175) and from external experts such as Matt Mullford (London School of Economics Methodology Institute)

Useful sources of advice

VFM Technical Notes available on MERLIN provide useful concise advice on diagnostic and analytical tools. The following Technical Notes are available and we will continue to update and expand these:

- Benchmarking
- Focus Groups
- Multi-Criteria Analysis
- Self Completion Questionnaires
- Bi-Variate Analysis
- Multi-Variate Analysis
- Using the Internet in VFM work
- Meta-Analysis
- Inter-Organisational Mapping

Summary. In selecting the right approach to use in a study the guiding principle should be the added value which it will deliver. The study methodology should never be an end in itself; its importance is the contribution it makes to our recommendations and in convincing our key audiences – Parliament and departments of the validity and authority of our findings and conclusions.

Who dares not stir by day must walk by night

King John Act I, Scene I

Technique	What is it all about?	When is it best used?	Is it more likely to be best used as a diagnostic or analytical tool?
1. Meta Evaluation/ Analysis	Synthesis of the findings from many evaluations, studies, and pieces of research to construct out an overall picture based on common and/or combined results and conclusions.	When there is a large amount of secondary evaluation available (eg health and social fields) where evaluation by study team may be prohibitive in terms of cost/time. Good for longitudinal analysis.	<ul style="list-style-type: none"> ● Meta evaluation at higher level is a good diagnostic tool to identify key questions/ issues. ● Meta analysis – more analytical/ statistical based on quantified data and drawing on combined results.
2. Organisational Mapping	Sets out in diagrammatic form the formal and informal relationships and links between and within organisations. Includes investigation of incentives and business processes to promote effective partnering.	To explain how organisations work together in pursuit of common goals (joined up government and service delivery).	Analytical to test strength/weakness of links eg common goals, lines of communication, sharing information, joint management meetings and partnership working.
3. Web Based Data Collection Tools	Using the web to identify and collect relevant material and to obtain stakeholder views on study questions. For example e.mail questionnaire, web-based survey, one to one internet interviews and web based focus groups.	At feasibility/ preliminary study to gather information on study area. During study to collect data and views.	<ul style="list-style-type: none"> ● Diagnostic tool for helping to select key questions for study. ● Can be used as a analytical tool if examining quality of Web sites.

Will evidence be more qualitative or quantitative?	Will it be costly to use? ¹	Will it take a long time to use? ²	Where do I go for further advice?
Meta Evaluation of overall conclusions more qualitative. Meta analysis of results more quantitative.	Medium – use of expert panel can reduce cost of surveying for relevant material.	Average	Dave Clark (Technical Note)
Qualitative – to identify perceptions and values as well as recognised formal links between organisations and partnerships.	Medium – may use other techniques to support analysis such as focus groups, structured interviews and survey.	Average	Robin Ryde/ Nick Lacy (Technical Note)
Both	Low	Below average	Nick Lacy

¹Costly to use: under £5k – low, £5k–£10k – medium, over £10k – high.

²Time to use: below average – under 4 weeks (ie could be completed during preliminary/feasibility stage), 4–8 weeks – average (ie easily completed during main study, over 8 weeks – high (likely to require piloting, data collection and analysis ie like a survey)

Technique	What is it all about?	When is it best used?	Is it more likely to be best used as a diagnostic or analytical tool?
4. Logic Models	Obtaining organisations opinions on the intended outcomes of programmes and services to map out the logical pathway from inputs through to outcomes and thus enable cause and effect to be better identified.	Where outcomes of activities may be uncertain and the factors which contribute to them unclear. Enables examiner to build up a good understanding of programme and expectations of customers and identify key performance points.	Analytical to identify either those activities which contribute directly to the outcomes of a project, programme or service and those results which are attributable to an activity.
5. Cost Benefit Effectiveness Analysis	Estimating the costs (direct and indirect) and expected benefits (tangible and intangible) from the outcomes of an activity.	Assessment of the overall benefits of recommendations and used in the assessment of options, priorities and decisions.	Analytical to determine the net worth of a proposed activity/decision.
6. Risk Assessment	Testing the quality of organisation's business risk management process to identify areas for examination, or evaluation of organisation's risk assessment. Good for identifying whether resources devoted to managing risk are consistent with level of risk to organisation and stakeholders.	At the problem formulation stage to identify key risks to achievement of organisation's objectives, risks to VFM and to outcomes. During study to consider quality of risk assessment for example for a major project, programme or public service. Will involve assessment of action is response to identified risk and contingency planning.	Both

Will evidence be more qualitative or quantitative?	Will it be costly to use?¹	Will it take a long time to use?²	Where do I go for further advice?
Qualitative	Low (much of the evidence collected via interviews).	Average	Nick Lacy
Costs – quantitative. Benefits – can be expressed in qualitative terms but usually given some quantitative measure.	Average – (depends on availability of primary data).	Below average (if computer database/software used to analyse data).	Nick Lacy
Both – impact of risk can be expressed in financial terms, in impact on reputation (high medium, low) or in terms of achievement of desired outcomes.	Medium	Average	Nick Lacy

¹Costly to use: under £5k – low, £5k–£10k – medium, over £10k – high.

²Time to use: below average – under 4 weeks (ie could be completed during preliminary/feasibility stage), 4–8 weeks – average (ie easily completed during main study, over 8 weeks – high (likely to require piloting, data collection and analysis ie like a survey))

Technique	What is it all about?	When is it best used?	Is it more likely to be best used as a diagnostic or analytical tool?
7. Cognitive Mapping	A versatile technique that can for example enable those directly involved (stakeholders) in service delivery to map out all potential factors contributing to a real or perceived outcomes.	At the start of a study to identify key questions/issues. Good for problem formulation.	Diagnostic
8. Multi-Criteria (Decision) Analysis	Maps out in a decision tree the options at key stages leading to expected and unexpected outcomes.	At the problem formulation stage to identify the key issues/ area the study could examine and the main stakeholders.	More a diagnostic tool but can be referred to throughout the study to test findings, conclusions and solutions.
9. Mystery Shopping	Covertly playing the role of a service user in order to experience and assess the service first hand.	For quality of service examinations both during diagnosis and analysis.	Both.
10. Multi-Variate Analysis	Allows the examiner to test for the key driver impacting on an observed output/outcomes – by assessing the effect of a change in one independent variable on the dependent variable.	Multi-variate techniques provide a robust way to make specific predictions about potential savings associated with changes in practice. It puts the examiner in a strong position to make soundly-based practical recommendations.	Analytical – heavy statistical input will require specialist advice.

Will evidence be more qualitative or quantitative?	Will it be costly to use? ¹	Will it take a long time to use? ²	Where do I go for further advice?
Quantitative	Low	Below average	Robin Ryde
Mainly quantitative but values/probabilities can be assigned the branches of the tree to identify most likely outcomes.	Average	Below average	Robin Ryde
Qualitative	Above average	Average	Robin Ryde
Quantitative	Above Average (but depends on size of sample/project)	Average	Robin Ryde (Technical Note)

¹Costly to use: under £5k – low, £5k–£10k – medium, over £10k – high.

²Time to use: below average – under 4 weeks (ie could be completed during preliminary/feasibility stage), 4–8 weeks – average (ie easily completed during main study, over 8 weeks – high (likely to require piloting, data collection and analysis ie like a survey)

Bibliography

- Modernising Government White Paper (CM 4310 1999).
- Professional Policy Making in the Twenty-first Century Cabinet Office September 1999.
- Wiring it up – Whitehall's Management of cross-cutting policies and services – a Performance and Innovation Unit Report – January 2000.
- Evaluation for the 21st Century Editors Eleanor Chelimsky and William R Shadish Sage (1997).
- Adding it up – Improving Analysis and Modelling in Central Government – a Performance and Innovation Unit Report – January 2000.
- Evaluation Research – An Introduction to Principles, Methods and Practice – Alan Clark with Ruth Dawson 1999 Sage.
- 'Comparing and contrasting auditing and evaluation: some notes on their relationship', *Evaluation Review*, 9 (4): 483 – 503 Chelimsky, E (1985).
- Practical Evaluation, Beverly Hills, CA: Sage. Patton, MQ (1982).
- Utilization-Focused Evaluation, 2nd edn, Newbury Park CA: Sage Patton, MQ (1986).
- How to use Qualitative Methods in Evaluation, Newbury Park, CA: Sage Patton, MQ (1987).
- Qualitative Evaluation and Research Methods, 2nd edn, Newbury Park, CA: Sage Patton, MQ (1990).
- A Measure for Measures: A Manifesto for an Empirical Sociology, London: Routledge, Pawson, R (1989).

- 'Evaluation methodology: back to basics', in G Mair (ed.), *Evaluating the Effectiveness of Community Penalties*, Aldershot: Avebury – Pawson, R (1997).
- Realistic Evaluation*, London: Sage – Pawson, R and Tilley, N (1997).
- 'Numbers and words: combining quantitative and qualitative methods in a single large-scale evaluation study', *Evaluation Review*, 9 (5): 627 – 43. Rossman, GB and Wilson, BL (1985).
- 'The methodology of evaluation', in RW Tyler, RM Gagne and M Scriven (eds), *Perspectives of Curriculum Evaluation*, Chicago: Rand McNally pp 39 – 83. Scriven, M (1967).
- 'Types of evaluation and types of evaluator', *Evaluation Practice*, 17 (2): 151 – 61. Scriven, M (1996).
- 'Evaluation in the political context', paper presented at the annual meeting of the American Psychological Association in Montreal. Weiss, CH (1973a).
- 'The many meanings of research utilization', M Bulmer et al. (eds), *Social Science and Social Policy*, London: Allen and Unwin. pp. 31 – 40. Weiss, CH (1986).
- 'Formative and summative evaluation: related issues in performance measurement', *Evaluation Practice*, 17 (2): 145 – 9. Wholey, JS (1996).
- Evaluation Research* Alan Clarke ISBN 0 7619 5095 8 1999.

List of credits

Front Cover	Portrait of Shakespeare by Soest.
Inside Front Cover	<i>Julius Caesar</i> , Royal Shakespeare Theatre, 1972 Mark Dignam as Caesar, photo by Joe Cocks.
Page 5	<i>The Tempest</i> , Royal Shakespeare Theatre, 1982 Derek Jacobi as Prospero, photo by Joe Cocks.
Page 17	<i>Richard II</i> , Royal Shakespeare Theatre, 1973 Ian Richardson as Richard and Richard Pasco as Bolingbroke, photo by Joe Cocks.
Page 41	<i>Macbeth</i> , Shakespeare Memorial Theatre, 1955 Laurence Olivier as Macbeth and Vivien Leigh as Lady Macbeth, photo by Angus McBean.
Page 61	<i>Hamlet</i> , Royal Shakespeare Theatre, 1992 Kenneth Branagh as Hamlet, photo by Malcolm Davies.
Inside Back Cover	<i>Measure for Measure</i> , Shakespeare Memorial Theatre, 1950 Maxine Audley as Mariana and Harry Andrews as the Duke, photo by Angus McBean.

© Copyright The Shakespeare Centre Library,
Stratford-upon-Avon.



*Our doubts are traitors,
And make us lose the good we oft might win
By fearing to attempt*

Measure for Measure Act I, Scene iv

NAO Website:
<http://www.open.gov.uk/nao/home.htm>

MODERNISING
GOVERNMENT

