

Unclassified

PUMA/PAC(99)1



Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

OLIS : 22-Jan-1999
Dist. : 26-Jan-1999

Or. Eng.

**PUBLIC MANAGEMENT SERVICE
PUBLIC MANAGEMENT COMMITTEE**

IMPROVING EVALUATION PRACTICES

Best Practice Guidelines for Evaluation and Background Paper

73742

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

PUMA/PAC(99)1
Unclassified

Or. Eng.

Complete document available on Internet: <http://www.oecd.org/puma/>

Disponible également en français sous le titre:

Vers de meilleures pratiques de l'évaluation

© OECD, 1999

Applications for permission to reproduce or translate
all or part of this material should be made to:
Head of Publications Service, OECD,
2, rue André-Pascal, 75775 Paris Cedex 16, France

TABLE OF CONTENTS

Foreword.....	4
Best Practice Guidelines for Evaluation.....	5
Improving Evaluation Practices: Background Paper	9
1. Introduction.....	9
2. Defining Evaluation	11
3. Managing Evaluations to Improve Their Use.....	23
4. Conclusions.....	39
References	42
Appendix: Reference Group on Evaluation.....	45

FOREWORD

A focus on results is a central element in recent public sector reforms in OECD countries. Evaluation is important in a results-oriented environment because it provides feedback on the efficiency, effectiveness and performance of public policies and can be critical to policy improvement and innovation. In essence, it contributes to accountable governance.

As a part of its work on Performance Management, the Public Management Service has studied evaluation in Member countries in order to identify key issues and practices to improve the use of evaluation. This report contains *Best Practices Guidelines for Evaluation* and a background paper, *Improving Evaluation Practices*.

The aim of the Guidelines is to identify key issues and practices that OECD Member countries should consider when seeking to improve the use of evaluation. They focus on management of evaluation activities in government and management of individual evaluations rather than on methodological questions. It must be emphasised that there is no single right way to organise and conduct evaluations. The choice of methods will depend on several factors, including the objectives of evaluation, the role of evaluation in a wider performance management framework, and institutional and political considerations. It is not the role of the Guidelines to determine when evaluation is the most appropriate input to the policy-making and performance management process. That decision will best be taken by the Member countries.

The Guidelines were reviewed and endorsed by both PUMA's Performance Management Network and the Public Management Committee.

This report and the Guidelines were prepared by Sigurdur Helgason of the OECD Public Management Service. The background paper was prepared by Marjukka Ala-Harja (Ministry of Finance, Finland) and Sigurdur Helgason. The Secretariat was assisted by a Reference Group of senior officials and experts from Australia, Canada, Sweden, the United States and the European Commission.

This report is published on the responsibility of the Secretary-General of the OECD. The views expressed are those of the authors and do not commit or necessarily reflect those of governments of OECD Member countries.

This report, along with other information about PUMA's work on performance management and evaluation, are available on PUMA's Internet site at: <http://www.oecd.org/puma/>.

Best Practice Guidelines For Evaluation

THE GUIDELINES

The Best Practice Guidelines for Evaluation are presented in three sections:

GETTING THE MOST FROM EVALUATIONS

This section defines evaluations, their objectives, main actors, and benefits and costs.

1. Definition and Objectives
2. Identify Main Participants
3. Assess Benefits and Costs

ORGANISING THE EVALUATION FRAMEWORK

This section discusses practices in relation to improving organisation and use of evaluations across the public sector.

4. Foster Evaluation Culture
5. Manage Evaluation Activities Strategically
6. Enhance Credibility

BUILDING EFFECTIVE EVALUATIONS

This section discusses practices and priorities for managing individual evaluations.

7. Ensure Links with Decision-Making Processes
8. Choose the Right Evaluator
9. Involve Stakeholders and Communicate Findings Openly

Getting the most from evaluations

Definition and objectives

- **Evaluations** are analytical assessments addressing results of public policies, organisations or programmes, that emphasise reliability and usefulness of findings. Their role is to improve information and reduce uncertainty; however, even evaluations based on rigorous methods rely significantly on judgement. A distinction can be made between ex-ante evaluations (or policy reviews) and ex-post evaluations. Many practices discussed in these Guidelines apply equally to both, even if their objectives are different.
- The main **objectives** of evaluations are to improve decision-making, resource allocation and accountability. This can be achieved through informing the public, informing key decision-making processes and encouraging ongoing organisational learning.
- Evaluations must be **part of a wider performance management** framework. They can supplement and improve it, but not replace it.

Identify main participants

- Successful evaluations are based on **collaboration between key participants** (evaluators, users and stakeholders), under the leadership of a “commissioner”.
- **Commissioners** are organisations that commission evaluations. The commissioner plans the evaluation, monitors its progress, receives the evaluation report, and makes decisions about further action. Commissioners may be ministries or central government agencies (e.g., the Ministry of Finance or independent evaluation and audit organisations). In some cases the commissioner may also be the evaluator.
- **Evaluators** are those organisations or individuals collecting and analysing data and judging the value of the evaluated subject.
- **Users** of evaluation may be policy-makers, the budget office, auditors, policy or programme managers and staff, users of services, etc.
- **Stakeholders** are those individuals or organisations that have an interest in the policy or programme being evaluated and the findings of the evaluation. Stakeholders and users are often the same actors.

Assess benefits and costs

- **Benefits of evaluations should outweigh their costs** and limitations. Both costs and benefits can be affected by careful management of evaluations and by choosing the appropriate evaluators and evaluation methods.
- The key value of evaluations is that they allow for **in-depth study of performance** and independent assessment of effectiveness of other performance management instruments. Potential benefits are the greatest for large policies or programmes.
- On the other hand, experience shows that evaluations have often been too **costly and time-consuming** compared to their real use and effect. There is also a risk of evaluations being used to slow the process of decision-making and justify inaction.

Organising the evaluation framework

Foster evaluation culture

- **Support** for evaluations is demonstrated through willingness of politicians, policy managers and central management agencies (e.g., Ministry of Finance), to make effective use of policy advice generated in evaluations.
- **Demand** for evaluation needs to be generated, specified and articulated by internal and external stakeholders. Evaluations without “ownership” by stakeholders are unlikely to have an effect. Institutional barriers to evaluation such as internal resistance can be reduced through consultation, aiming at creating mutual trust.
- The government can support an **evaluation culture** that encourages innovation and adaptation to a changing environment. The basic message should be that to stay relevant, organisations need to continue learning from feedback about results.
- **Training and professional dialogue**, competent evaluators, well-informed commissioners and enlightened and enthusiastic users all contribute to an evaluation culture.

Manage evaluation activities strategically

- **Organisation of evaluation should correspond to needs and priorities** in different policy areas. It may be appropriate to systematise and institutionalise evaluations in key policy areas where the costs of collecting data is high and information limited. However, a more flexible approach will often produce better results and prevent evaluations from becoming paperwork exercises. Special attention should be given to evaluation of activities that cut across many organisations.
- Central government agencies play an important role in managing the evaluation process; however, the **actual evaluations can be decentralised** to different actors at all levels of government.
- Development of evaluation skills in different organisations ensures the necessary **range of evaluation methods and perspectives** (e.g., drawing from both internal and external evaluators), and that each evaluation is designed in accordance with its unique set of issues related to objectives, focus, credibility and intended use.
- **Special funds** for financing evaluations can serve as an important incentive for evaluating public policies; however, they may also serve to encourage use of evaluation when other performance management approaches may be more appropriate.

Enhance credibility

- Lack of credibility undermines the use of evaluation findings. **Factors influencing credibility** include the competence and credibility of the evaluator, mutual trust between the evaluator and those evaluated, consultation and involvement of stakeholders and processes for communicating findings.
- Professional and ethical standards, and **methodological quality** of evaluation (encompassing issues such as relevant criteria, adequate data and evidence and reliable and clear findings) also have an effect on the credibility of evaluation. Quality assurance and open and frank dialogue can improve credibility by exposing and rectifying potential weaknesses in evaluations.

Building effective evaluations

Ensure links with decision-making processes

- Evaluation information can be an important factor in **policy formulation** to improve the quality of policy intervention and in the budget process to support priorities and savings. Relevant evaluations address issues that are significant for political, budgetary, management and other strategic reasons.
- **Objectives** of evaluation determine location, methodology and use of evaluation. The proposed use of evaluation should be clearly defined. Evaluations should be tailored to the characteristics of a policy intervention and evaluation methods should match the objectives of the evaluation, taking constraints such as costs and time into account. Building requirements for evaluations into policies from the start, and defining their objectives clearly, will improve the usefulness of evaluation and facilitate planning.
- **Planning** improves the management and quality of evaluation. The commissioner is responsible for planning evaluations, including defining objectives, criteria, data collection and methods. Timing is important, but the decision-making cycle is often unpredictable and decisions are often taken before evaluations have been finalised.

Choose the right evaluator

- **Self evaluation** by an organisation is appropriate when the main objectives are organisational learning and improved implementation. However, the time and skills of staff may be insufficient, the range of issues covered may be limited and the credibility of findings may also be questioned.
- Evaluation by **central management agencies** is appropriate when the objective is improving budget priorities and when it is important that the evaluator has close links with decision-making processes.
- Evaluation by **external evaluators** (e.g., research bodies and management consultants) is appropriate when the objective is to provide new perspectives on public policies or when there is a need for specialised evaluation skills. However, these evaluators may have limited understanding of the substance and the culture of the evaluated policy or organisation and offer theoretical evaluations.
- **Independent evaluation** is appropriate when the objectives are to improve accountability and transparency. However, policy managers, or the administration in general, may be reluctant to accept the findings and recommendations. Performance audits are often similar to evaluations. Their key features include independence of the auditor and a focus on accountability rather than improvement.

Involve stakeholders and communicate findings openly

- **Stakeholders**, including staff, can be appointed to evaluation commissions or involved through steering or advisory groups. Participatory evaluation methods can be used to create consensus and ownership for a change process. Dialogue with users and staff improves understanding and responsiveness to their needs and priorities. Participation must be managed due to the costs, time constraints and the risk of capture from such processes.
- **Presenting evaluation findings** openly increases credibility and creates pressure to act upon findings. Public availability of reports and meetings are useful to present and stimulate dialogue on findings. Judgements and recommendations based on clear criteria attract attention and promote subsequent action. Judgements should focus on overcoming problems rather than on assigning blame.

IMPROVING EVALUATION PRACTICES

1. INTRODUCTION

A focus on results is a central element in recent public sector reforms in the OECD countries. Evaluation is a tool for providing feedback on the results of organisations and programmes. This paper is mainly about programme evaluation, although many of the issues discussed may be equally relevant for other types of evaluations, e.g. evaluations of organisational performance, management reviews and performance audits.

Definition of Programme Evaluation
<p>Programme evaluation can be defined as: a systematic and analytical assessment addressing important aspects of a programme and its value and seeking reliability and usability of findings.</p>

Programme evaluation is pursued to improve decision-making -- to provide decision-makers with better information, and to enhance the accountability of public policies and programmes. It does not aim to resolve or replace judgements in making decisions, but it can ensure a degree of feedback about results of programmes that can be used to improve programme design and implementation.

When outlining the priorities for public sector reforms, the OECD Public Management Committee has pointed out that a well-performing public sector should continuously evaluate policy effectiveness. There are signs that interest in evaluation as a management tool is increasing. Compared with previous evaluation efforts there are now more realistic expectations, more widespread acceptance of less rigorous methodologies and greater understanding about utilisation in organisations. There is also a strong emphasis on more systematic, outcome-oriented evaluation with linkages to the budget process.

Making use of the findings of evaluations is not an easy task. The history of evaluation may be characterised as one of unfulfilled promises. Therefore, it is only natural that doubts are raised about its overall usefulness. Some see evaluation as a new management fad that creates bureaucracy but delivers few results. It is sometimes seen as inherently too theoretical to work in practice and having an effect on issues of marginal significance rather than on major policy choices. Others fear increased control or do not wish to be held accountable. Those primarily interested in the continuation of programmes may feel threatened by evaluation. Finally, questions can be raised about the value of evaluation in an environment where performance is continuously measured and even contracted for.

These arguments can be contested. Evaluation can fulfil an important role when properly used and integrated with the overall performance management framework. It can improve the efficiency and effectiveness of the public sector and hence strengthen the basis for public sector activities. The qualities of programme evaluation can complement and develop performance measurement, monitoring and auditing. By setting realistic expectations, confronting the problems of evaluation and using it in a pragmatic but purposeful way, governments can enhance their performance and accountability. The examples presented in this paper provide some guidance in this respect.

This paper discusses the role of programme evaluation in public management, and looks more closely at ways of improving its use in the public sector. The first part concentrates on defining evaluation. The second part presents issues worth considering when seeking to improve the use of evaluation.

2. DEFINING EVALUATION

2.1 CONCEPT

Evaluation persuades rather than convinces, argues rather than demonstrates, is credible rather than certain, is variably accepted rather than compelling (House, 1980, p. 73).

There is no general agreement on what evaluation is. The concept is defined in multiple, even contradictory, ways. This can be explained by the variety of disciplines (economics, policy and administration studies, statistics, sociology, psychology, etc.), institutions and practitioners active in evaluations and the wide range of issues, needs and clients that evaluation serves.

As evaluation is becoming increasingly recognised, virtually any type of feedback or inquiry may be called evaluation. A number of terms (review, follow-up, monitoring, audit, scrutiny, assessment) are used to refer to evaluative activities. Some experts have expressed concern about the possibility of the concept losing all its meaning. Furthermore, there are other feedback mechanisms, in addition to evaluation, that can be used to improve decision-making.

Not all feedback mechanisms constitute evaluation, although many of them have common characteristics. The difference between evaluation and other feedback mechanisms is shown in the following table.

Difference Between Evaluation and Other Feedback Mechanisms	
• Scientific studies	➔ Evaluations focus on practical use of information.
• Traditional audits	➔ Evaluations study public spending from a wider viewpoint - also questioning whether the objectives of the programme are appropriate and efficiently and effectively achieved (the distinction between audit and evaluation is often somewhat blurred and will be discussed later).
• Monitoring	➔ Evaluation is often conducted as a single effort and seeks to gain more in-depth information of a programme in question, although the existence of well-functioning, regular monitoring systems is often a necessary basis for successful evaluations.
• Performance measurement	➔ Evaluation strives for more: it tries to find explanations for observed outcomes and understand the logic of public intervention (however, well-designed performance measurement systems are, for example in the USA, often identified as one form of evaluation).
• Policy analysis	➔ Evaluation focus on <i>ex post</i> assessment. Policy analysis is sometimes defined as <i>ex ante</i> evaluation in studying future policy options.

The differing emphases are briefly presented in a following table. The dimensions reflect the different approaches to evaluation. Often there is a relation between the approach and the evaluation techniques applied. The differences are often linked to the process involved or the intended purpose of the evaluation.

Emphasis of Evaluation	
• Timing	→ <i>Ex ante</i> , concurrent/intermediate, <i>ex post</i> or a combination of these.
• Flexibility of framework	→ Assessment against the objectives of the programme or assessment involving also the appropriateness of objectives. Assessment of intended and perhaps of unintended effects.
• Judgement or presentation	→ Judging the success or failure of a programme - even suggesting improvements or alternatives - or presenting the findings and leaving the conclusions to decision-makers.
• Location and autonomy	→ Internal (e.g. self-evaluation) or external assessment or a combination of these (e.g. semi-independent government evaluation units).
• Expertise or negotiation	→ Assessment by evaluation experts applying scientific and technical procedures to secure the objectivity of results, or a partnership where aims and methods -- and even findings and recommendations -- are negotiated with stakeholders to maximise the use of findings and secure democratic participation.
• Use	→ Gaining more insight into a programme or deciding its value. Helping into fine-tune a programme or questioning its rational and proposing more fundamental changes.

By emphasising different aspects, evaluations can be classified into various types. In practice, these types of evaluations are often combined, as can be demonstrated with the case of summative and formative evaluations.

Types of Evaluations	
• Summative evaluations	→ are often carried out when the programme has been in place for some time (<i>ex post</i> evaluation) to study its effectiveness and judge its overall value. These evaluations are typically used to assist in allocating resources or enhancing public accountability. The clients are usually external, such as politicians and other decision-makers. The objectivity and overall reliability of findings is considered important, and external evaluators are therefore often commissioned to conduct the evaluation. Questions of outcome and overall relevance of the programme are expected to be addressed.
• Formative evaluations	→ are usually undertaken during the implementation of the programme (intermediate evaluation) to gain further insight and contribute to a learning process. The purpose is to support and improve the management, implementation and development of the programme. The evaluators as well as clients are typically internal, usually programme managers (self-assessment). The objectivity of findings is often not the main concern: more emphasis is put on the direct applicability of results. Operational questions, monitoring of events and to some extent impacts are expected to be addressed.

Programme Evaluation

Evaluation can study organisations, functions, procedures, policies, projects, etc. This paper concentrates on the assessment of public programmes, i.e. programme evaluation. “Programme” usually refers to a set of organised activities with a specific time schedule and budget to put policies into action, i.e. to create conditions to achieve the desired political goals. The term can refer to relatively narrow activities or to

major cross-cutting policy programmes. However, it should be pointed out that the “programme” concept is understood in different ways in different countries and in some cases not used at all.

As in the case of evaluation there is no single definition of programme evaluation. At its simplest, it includes assessment of a programme’s achievement against its objectives. It is increasingly suggested that programme evaluation can encompass any stage in a programme’s life-cycle. In addition to *ex post* assessment this includes *ex ante* evaluation and concurrent evaluation (which is very common in practice). The former type of prospective analysis is also often called policy analysis or appraisal. While policy analysis explores policy options and likely effects, evaluation studies actual effects and judges the value of policies. Nevertheless, the concepts are inter-linked and the assessment processes are interactive.

Examples of Definitions of Programme Evaluation
<ul style="list-style-type: none"> • The key issues in program evaluation are to attempt to determine what the effects of the program have been over a specific period of time and to determine the extent to which the specific program has caused these effects. (Hatry, 1980). • Program evaluation is the application of systematic research methods to the assessment of program design, implementation and effectiveness. (Chelimsky, 1985). • Programme evaluation is the systematic assessment of a programme, or part of it, to firstly assist managers and other decision makers to assess the appropriateness, efficiency and effectiveness of a programme, secondly to assess the case for establishment of new programmes, or extensions to existing programmes and thirdly to decide whether the resources for the programme should continue at current levels, be increased, reduced or discontinued. (<i>Doing Evaluations, A Practical Guide</i>. Australian Ministry of Finance, 1994). • A programme evaluation is commissioned by an organisation with the aim of determining how well a programme is working, and whether it needs to be improved or abandoned. (Mayne, 1994). • Evaluations should normally have elements of being analytical, systematic, reliable, issue-oriented and user-driven (Evaluating EU Expenditure Programmes. A Guide. European Commission, 1997).

Despite the various definitions, some characteristics are common to all programme evaluations:

Programme evaluations are systematic, analytical assessments, addressing important aspects of a programme and its value, and seeking reliability and usability of findings.

The lack of a consensus on the concept of evaluation can cause confusion, but it can also be seen as a fruitful, practical point of departure reflecting the various implicit paradigms in social sciences. However, from the point of view of this paper the question of whether one should apply a broad or narrow definition of evaluation is not particularly important. The objective is to identify good practices that would apply to different definitions.

2.2 OBJECTIVES

The objective of evaluation is not to seek absolute truth but to provide insight and well-justified views on policy programmes. Evaluation is not a substitute for political decision-making, but it enables decisions to be taken on a more informed basis.

Management is often described as a cycle of planning/preparation, budgeting, implementation and evaluation. Evaluation is the link which closes the circle and completes the feedback loop. The evaluated policy or programme is improved, expanded or replaced.

The principal aims of evaluation may be characterised as improving decision-making, resource allocation and accountability.

Objectives of Evaluation	
<ul style="list-style-type: none"> • Improving decision-making 	<p>➔ Evaluation provides information on the impacts of policies. It can be seen as a step from formal control mechanisms to steering through information, from transactions control to strategic control, from input management to management by outcomes. It assists policy-makers in judging the value of public programmes, in improving or challenging them and in designing future programmes. Evaluation also contributes to the learning process of those managing and implementing programmes by increasing their understanding of the work they carry out and helping them to develop it. In addition, other interests (e.g. scientific) may be satisfied in gaining more insight to the functioning of society.</p>
<ul style="list-style-type: none"> • Assisting in resource allocation 	<p>➔ Most OECD Member countries have been facing budgetary pressures. Evaluation can help policy-makers and budget officials improve resource allocation and make better use of existing funds. They assist in finding more effective and efficient ways of reaching the desired results and shifting expenditure to facilitate the development of new priority areas - and contribute to finding the right level of funding for programmes or even prioritising them when budget cuts are needed. Evaluation has been closely linked to innovations in the budgetary process, for example in Sweden and Australia.</p>
<ul style="list-style-type: none"> • Enhancing accountability 	<p>➔ Evaluation can contribute to the accountability of public policies, which is felt to be increasingly important in OECD Member countries. This is due to the increased autonomy of the administration caused by major management reforms and concerns among citizens in some countries about the legitimacy of the public sector. Evaluation can improve transparency and accountability to the parliament and public at large by shedding light on the impact of government activities. Hence, it can strengthen the legitimacy of ongoing programmes and secure the sustainability of results. It also helps citizens to learn what is being achieved with their money and consider alternatives to implemented policies.</p>

In addition, there are some factors that make evaluation especially timely:

Factors that make evaluation timely	
<ul style="list-style-type: none"> • Public sector reforms, strategic management 	<p>→ Evaluation is one of the tools highlighted in recent public sector reforms. It is often presented in the context of developing performance measurement and strategic management. It is argued that assessing performance, especially self-evaluation, should become a normal part of public management. There is also a growing need to assess ongoing public sector reforms in OECD countries in order to exploit their potential, develop them further, sustain achieved results and strengthen their public support.</p>
<ul style="list-style-type: none"> • Institutional requirements 	<p>→ Institutions that finance or co-finance public programmes often require recipients to carry out evaluations or conduct evaluations themselves (e.g. the Structural Funds of the European Union and the World Bank). This is also the case when governments increasingly institutionalise their evaluation activities and require ministries and agencies to assess their results.</p>

Organisations view evaluation in various ways and use it for different purposes. Different objectives lead to different approaches to evaluation. Yet, **on a more general level the goal of evaluation may be defined as organisational learning.**

Examples of stated evaluation purposes
<ul style="list-style-type: none"> • Evaluation can provide credible, timely and objective findings, conclusions and recommendations to aid in resource allocation, program improvement and program accountability (Australia). • ... produce sufficient, appropriate and credible information which will help to decide whether EU expenditure programme should be continued, unaltered, changed or discontinued and whether the corresponding budgetary appropriations should be increased or reduced (The European Commission, DG XIX, Assessment checklist). • Evaluation is seen as adding value by focusing on what is really working, what is not, and by finding innovative ways of achieving government goals more cost-effectively (Canada, <i>In Search of Results: Performance Management Practices</i>).

2.3 DEVELOPMENT

In some OECD countries evaluation has a formal role in the policy-making process, while others are just beginning to reflect on the possibility of systematising evaluation activities. Countries such as the United States, Sweden and Canada have a long history of evaluation with a firmly institutionalised evaluation system. Australia and the Netherlands have also developed sophisticated evaluation systems.

The maturity of evaluation practices also varies among policy sectors. Evaluation has a long tradition in a number of countries in areas such as developmental aid, labour market policies, health and education.

Historical phases

Evaluation activities in national administrations have developed in several phases. Maturity can be characterised by the length of experience, the regularity of activities and the degree of institutionalisation. Three stages of development have been distinguished (Derlien, 1990). (Although assessments of public programmes were conducted before the first phase, for example in Sweden).

The first wave of evaluation in the 1960s and 1970s was largely linked to social-liberal governments launching public programmes to solve social problems. Planning and programming processes were introduced to improve policy-making. Favourable fiscal conditions and the increased status and supply of social science knowledge contributed to this effort. The focus was on improving programmes, and programme managers were interested in using evaluation as a feedback mechanism.

The second wave in the 1980s -- although less striking and rapid -- was stimulated by predominantly conservative governments attempting to curb public programmes given fiscal constraints. Evaluation was thought to be useful in reconsidering the justification of policies and rationalising resource allocation within the budget. Ministries of Finance and Audit Offices were active in developing evaluation.

Since then evaluation has increasingly become the concern of parliaments. Problems with public sector legitimacy and the need to strengthen the accountability of government call for new, more sophisticated control mechanisms such as evaluation. The preliminary findings of an ongoing cross-national study (Gray, 1997) suggest that developments in new public management have many implications for the function and practice of evaluation: evaluation is substituted (new internal markets are intended as intrinsic evaluators), complemented (the regulation of privatised monopolies is creating a huge demand for evaluation) and offered new subjects (evaluating new tools, for example internal contracts).

A firmly established evaluation tradition is inclined also to see the need for evaluation of evaluation. This meta-evaluation work is probably most advanced in the USA, where the amount of evaluation makes it possible to draw some general conclusions. Yet, meta-evaluations are still fairly rare and not much is really known, for example about the use of evaluation.

Evaluation and audit

To clarify the role of evaluation and its current development, it is useful to study its relation to traditional auditing. Evaluation and audit are historically separate functions with separate institutions. The respective professional cultures of evaluators and auditors differ in various ways, for example in terms of recruitment and socialisation. The objectives are different: audits are traditionally used to ensure financial regularity and accountability for resources while evaluation aims to promote economy, efficiency and effectiveness.

Audit has a long tradition of serving the parliamentary oversight function in many countries. It underlines legal compliance and accountability of public organisations. Traditionally, auditing seeks evidence that established rules, regulations, procedures or mandates have been followed. The independence of audit institutions enables them to submit credible opinions and proposals.

Evaluation has not as audit, established itself as a regular, institutionalised activity with a clear concept and agreed objectives to the same extent. Evaluation is typically more participatory and less independent than audit and accepts the views of stakeholders to a greater extent. Evaluation is used more as a tool for programme management than for public accountability.

Despite the obvious differences between traditional audit and evaluation, a question has been raised whether there is a need to bring these functions closer together or whether such convergence is already taking place. Boundaries have begun to blur: audit has extended its scope into questions of performance, and traditional financial audits have been supplemented with performance audits. Performance audits, carried out for example in Sweden, the United Kingdom (value-for-money-audits) and the Netherlands, are an example of assessment similar to programme evaluation. The methods and criteria applied are very

similar. Common networks, such as the European Evaluation Society (EES), also indicate convergence of professional cultures.

However, contradictory signs exist as well. The convergence of evaluation and auditing is restrained by the multiple and contested identity of evaluation, its social-science-based culture and political resistance to more evaluation by audit institutions. Financial audit may also become more topical, with more reliance on outsourcing and innovations such as accrual accounting. It has been suggested that audit institutions should not aim to become primarily evaluation units or management consultants, but instead clarify their current roles and responsibilities as independent guardians of the probity, legality and efficiency in public spending (Pollitt & Summa, 1997).

Institutionalised evaluation

Some examples of evaluation practices are presented below to illustrate the development of evaluation.

In Sweden...
<p>Evaluations were first undertaken in the 1950s or even earlier by public commissions preparing policy decisions. The commission system combined the views of stakeholders and introduced research findings into decision-making. From the 1960s evaluation began to be viewed as an activity that could continuously provide decision-makers with information. Specialised research bodies and agencies were founded with evaluation as their main task.</p> <p>In the beginning of the 1980s there arose a need to improve decision-making by systematising evaluation activities and linking them more firmly to the budget process. The evaluation system as linked to the new budget process has gone through several fairly pragmatic adjustments since its creation. At present annual reports of agencies are the main source of follow-up information submitted by agencies to ministries. In-depth evaluations of agencies' activities (which were earlier to be carried out every third year) are conducted only upon special request by a ministry.</p>

In Australia...
<p>Evaluation became an integral part of Australia's public management reform process in the mid-1980s when the government's evaluation strategy was launched. Evaluation has been systematically integrated into corporate and programme management and planning. All public programmes, or significant parts thereof, are reviewed once every three to five years; all major new policy proposals include an evaluation strategy, ministries are currently required to provide an annual evaluation plan; and results of major evaluations are expected to be made public.</p> <p>Recently the continued usefulness of the mandatory requirements has been reassessed. It is likely that the requirements-based approach will be replaced by an integrated performance management framework, encompassing performance monitoring and evaluation, based on good practice principles. The shift recognises that evaluation has become more commonplace and widely accepted, and is intended to integrate performance management activity directly with ministries' strategic and business planning. It also recognises that more attention is being paid to the quality, utilisation and effectiveness of evaluations.</p> <p>The responsibility of planning, carrying out and using evaluations lies with individual agencies and departments. The Department of Finance is the co-ordinator of the evaluation strategy, providing training, encouragement, guidance, quality assurance and sometimes participating in selected major evaluations. The Auditor General may review the operation of evaluation activity.</p>

In the Netherlands...

The Court of Audit recommended in the early 1990s that ministries pay more attention to programme evaluation. The government established a three-track strategy: first, the amendment of regulations setting out the responsibilities of ministries for evaluation; second, the preparation of annual reports to accompany ministerial budgets to provide information on current and future evaluation activities, findings and their utilisation; and third, the management and co-ordination of evaluation at ministerial level. Government also pointed out that evaluation should be not only external but also internal and not only *ex post* but also *ex ante* and intermediate.

Later, government outlined evaluation measures and responsibilities. Policy directorates are responsible for evaluation themselves: they can adopt either a centralised or decentralised approach in managing activities. Internal budget directorates are responsible for co-ordinating, promoting and ascertaining the use of evaluation and ensuring that necessary advice, guidance and research expertise are provided. They draw up evaluation programmes for individual projects, encourage the periodic evaluation of policies and monitor the quality of evaluation and its use. The heads of ministries are ultimately responsible for evaluation. Other directorates provide help with their special expertise (on personnel and organisation, auditing, legislation). The Ministry of Finance has the same responsibilities for promoting and co-ordinating as internal budget directorates. The Court of Audit reviews the efficiency of evaluation management and organisation and publishes reports on them.

In Canada...

Evaluation has been part of the management of departments since 1977. Review activities are firmly institutionalised. Departments are responsible for ensuring that their programmes are evaluated. The present review policy requires departments to conduct internal audits according to established standards in areas of significance or risk; to conduct evaluations of key policies and programmes according to established standards; to establish performance monitoring practices, including client-oriented service standards, to ensure that departments and managers have and use credible information on key aspects of programme performance; to conduct other types of reviews such as policy, regulatory, programme and operational reviews, client surveys and special studies; and to use internal audit, evaluation, performance monitoring and other reviews in a co-ordinated and complementary fashion.

Evaluations are often carried out by specialised corporate evaluation units. The Treasury Board Secretariat is responsible for the policy, co-ordination and monitoring of evaluation. The Office of the Auditor-General has a major role in reviewing performance in the public sector and reporting on the adequacy with which evaluation responsibilities are carried out.

In Finland...

As in Sweden, there is a tradition of appointed commissions preparing political decisions. Currently the use of evaluation is well established in some sectors (development aid and higher education), although in general evaluations are conducted on an ad hoc basis. However, according to a recent survey there is a great deal of interest in evaluation in the administration. To obtain more systematic and profound information on the recent public sector reforms, a major evaluation programme consisting of 12 individual studies was carried out in 1995-1997. Another aim of this project was to find ways to strengthen the evaluation culture in the administration. Training courses are being organised, a handbook has been compiled and an attempt is being made to encourage ministries to report on their evaluation programmes and their results in their annual reports.

In the United States...

Evaluation has a long and extensive tradition in the United States. The federal government as well as some state and local governments have evaluation units. At the federal level these are usually within individual departments. The General Accounting Office (GAO), which assists the Congress in its oversight of the executive branch and encourages effective management and accountability, undertakes a large number of performance audits and programme evaluations. Agencies responsible for implementing programmes also conduct their own programme evaluations. Agencies have to include in their strategic plan a description of programme evaluations used and a schedule for future evaluations. In addition, the Congress has frequently inserted requirements for individual, usually large-scale, evaluations into legislation.

In the European Commission...

The European Commission has for some time had institutionalised evaluation practices in certain central policy areas of Community policy, such as development aid, research and technology policy and programmes financed through the Structural Funds. In the context of the Sound and Efficient Management initiative of the Commission (SEM 2000), organising evaluation of EU expenditure on a more systematic, timely and rigorous basis across Community policies has become a central objective. It is stated that the operational services are responsible for ensuring the evaluation of their programmes. An official, sector or unit has to be designated as responsible.

Every service must design its annual evaluation plan, the first of which were presented in the end of 1996. These plans are compiled, taken note of by the Commission and circulated as its Annual Evaluation Programme. Evaluation reports produced by the operational services of the Commission serve as a basis for an Annual Evaluation Review, which is submitted to the Commission with the Annual Evaluation Programme by the end of January each year. In addition, the operational services are expected to submit an *ex ante* evaluation report every time they propose an action incurring community expenditure. They should also ensure when preparing new programmes that evaluation is planned to fit into the process.

It is recommended that actions funded through the annual budget should be evaluated at least once every six years. In addition to operational services, DG XIX (Budget) and DG XX (Financial control) carry out various tasks supporting evaluation activities. To further support the evaluation culture a network of evaluators has been established, a manual setting quality norms has been compiled (European Commission, 1997) and training seminars, demonstration projects and a help desk are being organised.

2.4 PERFORMANCE MANAGEMENT AND EVALUATION

Many governments have introduced comprehensive approaches for managing performance of government organisations. Evaluation is clearly a performance-oriented instrument and should therefore be seen as a part of a wider performance management framework. However, the relations between evaluation and performance management are not always clear.

Evaluation and performance management...

- Evaluation was well established in many governments before the introduction of performance management, and new approaches did not necessarily incorporate evaluation.
- New performance management approaches were developed partly in response to perceived failures of evaluation, i.e. that the actual use of evaluation findings was limited compared to the resources put into evaluations.
- Different organisations may be responsible for evaluation on the one hand and performance management on the other.
- Performance management involves reforming core management processes and is thus the responsibility of managers within each organisation, whereas evaluation has been seen as a specialised method, often carried out by external experts or evaluation units.

Failure to consider and clarify relations between evaluation and performance management can lead to waste, duplication, conflicting signals about aims and results, and tensions among organisations and professional groups. Most governments see evaluation as a part of the overall performance management framework, but the degree of integration will depend on the role of evaluation. There seem to be three main approaches (see the box below).

Approaches to the integration of evaluation and performance management
1. Evaluation is seen as a separate and independent function, with clear roles <i>vis-à-vis</i> performance management.
2. Evaluation is seen a separate function but integrated with performance management.
3. Evaluation is not seen as a separate function but integrated with individual performance management instruments.

The first approach sees evaluation as relatively independent from performance management, even if it relies on performance measurement and other information produced through performance management. The basic perspective is that performance management is like any other management process within an organisation that has to be subjected to evaluation. Too much integration of evaluation and performance management would undermine the capability of evaluation to assess the quality of performance management and to make useful suggestions on how to improve it. Integration would also undermine the necessary independence of the evaluation function.

The second approach puts less emphasis on independence but sees evaluation as one of many instruments used in the overall performance management framework. Evaluation can be superior to other instruments in a number of respects. It allows for more in-depth study of performance than performance measurement. It can cover a longer period; incorporate factors that are difficult or expensive to assess through ongoing measurement, such as outcome or impact assessment; and analyse causes and effects in more detail, bringing performance into its proper context. Evaluation can be used to give a comprehensive image of the performance of a programme or organisation, as opposed to the somewhat fragmented picture given by other approaches. In this approach, evaluation is organised as a separate function to ensure expertise and is used selectively and periodically to produce key information not produced through other processes. The cost of conducting evaluations means that they will have to be limited to the most important programmes or organisations.

The third approach emphasises the benefits of incorporating evaluative methods into individual performance management instruments. There are many examples of this. Performance information can be analysed, interpreted and evaluated; extensive benchmarking studies often involve considerable evaluation of factors that explain differences in performance; evaluation is one of the key elements of strategic management (e.g. evaluating key factors in the environment, SWOT analyses); performance contracts are evaluated periodically; and quality management approaches often involve self-assessment. The basic perspective here is that evaluation should be an ongoing activity, integrated into key management processes.

These three approaches are not mutually exclusive, i.e. a government can use more than one approach. The second approach has clearly been gaining momentum. Despite this, some degree of independent evaluation capacity is being preserved, notably in performance audits carried out by audit offices. There is growing awareness about the benefits of incorporating evaluative methods into key management processes. However, most governments see this as supplementing rather than replacing more specialised evaluations.

2.5 METHODOLOGICAL ISSUES

Although this paper concentrates on the use of evaluation, it is relevant to briefly address some of the major methodological approaches and problems related to evaluation. Since the methodology and the use of evaluation are interlinked, these questions will come up again in the later discussion.

Approaches to evaluation

There are various approaches to evaluation incorporating assumptions about the nature of evaluation knowledge, how new reliable knowledge can be created and to what use it can legitimately be put. Furthermore, there often exists some relationship between the approach and the tools and techniques chosen for evaluation. Four main approaches and some evaluation methods usually related to them are presented in the following table (Pollitt, 1996).

Approaches and Methods	
<ul style="list-style-type: none"> • Experimental evaluation 	<p>→ Incorporates the positivistic idea of applying the methodology of natural sciences to the engineering of public programmes. In experimental designs, randomly assigned groups that are perceived as similar are compared after applying a programme to one of them. Attributing the observed effect to the programme depends on the assumption that there are no other rival causes for the outcome.</p>
<ul style="list-style-type: none"> • Pragmatic evaluation 	<p>→ Focuses above all on the usefulness of the evaluation findings. Evaluation is to be geared to the objectives and working practices of key decision-makers. Political feasibility, timeliness and cost are important. Quasi-experimental designs (where alternative explanations to outcomes are pursued to be ruled out by means other than random assignment), case studies and various descriptive methods are typically used as evaluation tools.</p>
<ul style="list-style-type: none"> • Economic evaluation 	<p>→ Has often been conducted slightly apart from evaluation in general. Tools include, for example, cost-benefit analysis and cost-effectiveness evaluations. They introduce the information on costs of the programme as a criteria in evaluations.</p>
<ul style="list-style-type: none"> • Naturalistic evaluation 	<p>→ Rejects both experimental and economic evaluation, claiming that society is socially constructed and constantly changed by the interaction of individuals. Therefore evaluation cannot provide right, objective answers but instead can act as facilitator to produce consensus among the stakeholders. Methods applied, such as ethnography, participant observation, case studies, etc., aim to describe and clarify the values and assumptions of the stakeholders and support their negotiation in the evaluation process.</p>

There is no best way to carry out evaluation. Various approaches and methodologies encompass different success factors. The role of evaluation varies in different learning and decision-making processes (e.g. economic and experimental evaluation used for budgetary decision-making, and naturalistic evaluation used for programme improvement and design). The important issue is to design the evaluation to suit the evaluated subject, desired evaluation information, available resources and the use of findings.

Methodological problems

Methodological problems are intrinsic in all approaches to evaluation. For instance, problems related to causality are common to social sciences in general. Conclusive evidence of cause-effect relationships can rarely be established, since controlling all relevant variables is seldom possible. Experimental evaluation design is often difficult, expensive and lengthy, if not impossible to apply in practice. Even if experimental evaluation design is used, generalisation beyond the conditions of the experiment is usually uncertain. However, in many evaluations, for example when evaluating the implementation or management of a programme, establishing a firm causal relationship is not the major objective of evaluation.

Choosing criteria for evaluation may be problematic simply because the intended objectives of public programmes are often multiple, vague, hidden, evolving and even conflicting. Another difficulty is deciding whether to focus only on how official objectives are achieved or to take a broader view and study all the effects of the programme. The latter gives a more comprehensive picture of the outcome of the programme but is more complex and time-consuming.

Setting an appropriate time period over which the programme is evaluated is difficult but critical, as relevant outcomes should have sufficient time to mature. However, the information's usefulness may diminish if the evaluated programme is changed before the evaluation is finalised or evaluation findings applied.

Choosing methods for evaluation is also problematic. Various ways of collecting and analysing data provide different perspectives on the evaluated programme. There is, for example, a considerable discussion on the advantages and disadvantages of using quantitative and qualitative methods in evaluation.

It is often difficult to identify programme outcomes. As pointed out earlier, there may be various reasons (other than the programme) that wholly or partly generate the outcomes. Thus, it may be difficult to determine what evidence is sufficient to draw valid and reliable conclusions about outcomes. The ability to gather evidence is often restrained by available resources. Another problem concerning conclusions is whether evaluation findings can be generalised to other times, places or subjects. This is of particular importance when evaluation is expected to contribute to future policy decisions.

Methodological problems in evaluation (only some of them have been covered here) can be dealt with when limitations are recognised and problems properly addressed. This requires specific knowledge and skills that can be gained by training staff and commissioning external expertise to conduct evaluations. Also, combining different methods is often the most fruitful approach to lessening methodological problems in evaluation. Furthermore, appropriate quality control mechanisms (see later discussion on ensuring technical quality) can be set up to guide the evaluation process.

Even if evaluation cannot provide definitive answers, it can add useful information to the discussion about performance: some credible information on what happened and why is better than none.

3. MANAGING EVALUATIONS TO IMPROVE THEIR USE

In addition to methodological problems, the other major issue has been the use of evaluation. Evaluation is not knowledge for its own sake, but knowledge for action. It is important that evaluation material be used in decision-making. If not, the administration will soon lose interest and evaluation will become an empty exercise.

It has been stated that for evaluation to have an impact on decision-making both a good quality supply and enthusiastic and sincere demand are needed (Mayne 1994; Mayne, Divorski, Lemaire, 1998; Boyle, 1997). This chapter will discuss how to advance these two aspects of the evaluation “market”: how to create opportunities for using evaluation.

It is difficult to define what constitutes the use of evaluation. One definition of use of evaluation is to make a distinction between programme and organisational use. The former relates to individual evaluation studies and the latter to the cumulative (learning) effect of series of evaluations on an organisation. Programme use can be further divided into instrumental use, where evaluation supports a specific action (confirmation, modification or termination) and conceptual use, where no immediate action is taken but evaluation findings are considered and information is acquired to be used later. (Mayne, 1994).

Use of Programme Evaluation
<ul style="list-style-type: none"> • Major policy changes • Programme improvements (organisational learning) <ul style="list-style-type: none"> • Improved budget allocation • Support for accountability

In any case, use of evaluation findings is not easy to demonstrate. Potential users have to be identified and the period of expected use defined. Use can involve concrete improvement but can also be less concrete, such as greater insight into a programme, organisational learning or accountability.

It is also possible to distinguish between the use of evaluation and its impact. The question here is how much impact is required to signify actual use of the evaluation findings. Experience suggests that evaluation is used more for improving existing programmes than for supporting fundamental changes in policies and resource allocation. This has been the case in Sweden -- although it should be pointed out that major policy changes may be supported by cumulative evaluation information gathered over many years. Also, many individual evaluations may not suggest substantial changes to programmes.

All these factors make assessing the use of evaluations a complex issue. Although there are not many systematic studies on use, some examples are provided in the concluding chapter.

Because evaluation is practised in a political context, the problem of ascertaining use will never be fully resolved. Evaluation is only one source of information and has to compete with other factors influencing

decision-making. It is almost impossible to find out to what extent evaluation findings have influenced decisions. Evaluation may have the greatest impact on policy measures of low political priority or when there is a broad agreement on the general direction of policy. Moreover, politicians emphasise good intentions when deciding to make public interventions, but they may not always be interested in studying their effects. Behind the reluctance to commission evaluations and discuss their findings may also lie a fear of shortcomings being exposed. Politicians may not always want to formulate policy goals precisely, which makes evaluation difficult. The fact remains that decisions will and have to be made regardless of the quality of the material feeding into the decision-making process. Bearing this in mind, it is possible to foster the use of evaluation by open consultation with decision-makers.

Performance-oriented culture that values evaluation does not emerge overnight. However, some concrete steps can be taken to make evaluation count and to enhance its use by the administration, decision-makers and the public at large. The following section will highlight some of the practices that can maximise the use and value of evaluation. However, it should be pointed out that different ways of enhancing the use of evaluation may apply in different contexts.

The first three issues focus on creating an effective demand and an environment conducive to carrying out programme evaluation. The next three issues concentrate on developing an effective framework for conducting evaluation. The following nine issues focus on practical aspects of the evaluation process itself -- how it can be managed to promote the use of its findings. The final issue addresses training and support as a vital factor in creating an evaluation culture in the administration.

Issues in Improving Evaluation Practices	
• <i>Issue No. 1:</i>	Gaining Support From the Top
• <i>Issue No. 2:</i>	Generating Effective Demand
• <i>Issue No. 3:</i>	Setting Realistic Expectations
• <i>Issue No. 4:</i>	Systematising Evaluation Activities
• <i>Issue No. 5:</i>	Linking with the Budget Process
• <i>Issue No. 6:</i>	Choosing the Appropriate Evaluator
• <i>Issue No. 7:</i>	Planning Evaluations
• <i>Issue No. 8:</i>	Timing Evaluations Appropriately
• <i>Issue No. 9:</i>	Meeting User Needs
• <i>Issue No. 10:</i>	Ensuring Relevance
• <i>Issue No. 11:</i>	Involving Stakeholders
• <i>Issue No. 12:</i>	Ensuring Methodological Quality
• <i>Issue No. 13:</i>	Making Judgements and Recommendations
• <i>Issue No. 14:</i>	Communicating Findings
• <i>Issue No. 15:</i>	Monitoring or Following-up
• <i>Issue No. 16:</i>	Recognising Needs of Staff for Training and Support

Issue No. 1:
GAINING SUPPORT FROM THE TOP

It is important that managers and politicians understand the value of evaluation and have realistic expectations about its use. This can be expressed through adoption of a formal evaluation strategy for the government. The role of the Ministry of Finance and other central agencies is often essential in ensuring support and credibility. Managers should be involved in the evaluation process, providing inputs as

required. When the objective is mainly to improve the implementation of an existing programme, the support of programme managers is essential.

In the Netherlands...
In the early 1990s, the government established a strategy stating the functions of evaluation and strategies for promoting it, and defining necessary initiatives and recommendations. Although policy directorates are responsible for organising evaluations, the heads of ministries are ultimately responsible. The Ministry of Finance and internal budget directorates are responsible for co-ordinating, promoting and ensuring use and making sure that necessary advice, guidance and research expertise is provided.

Issue No. 2: **GENERATING EFFECTIVE DEMAND**

It has been suggested that one reason behind the inadequate use of programme evaluation is the supply-driven nature of the market. Generic demand for evaluation may need strengthening. Demand needs to be generated and clearly articulated if evaluation is to be of practical use and lead to improvements in programmes (Mayne, Divorski, Lemaire, 1998). Customers have to specify their evaluation interests and make them known. There are several incentives that can be used to foster demand. These have been described as *sticks*, *carrots* and *sermons* (Boyle, 1997; Toulemonde, 1998).

Incentives	
• Sticks	→ Sticks can be used to ensure that evaluations are undertaken and relevant questions asked. This can be done by making evaluations mandatory or giving an external stakeholder the right to ask evaluation questions. The quality of evaluations can also be controlled by specific mechanisms, i.e. steering or review groups (France, Canada). Sometimes increased competition between government units encourages them to provide more evaluation information on their activities and results.
• Carrots	→ Carrots are used to develop demand. This can be done by earmarking funds for evaluation and by giving recognition to evaluation efforts (publicity, prizes, promotions, improvements in content of work, performance-related pay). Extra work caused by evaluation can be balanced by reducing input and process controls.
• Sermons	→ Sermons can be given to underline the role of evaluation and to promote a culture that values its use in decision-making. This message can be sent through conferences, training courses and visits, newsletters and journals, peer group networks, awards, etc. Most importantly, acting on evaluation findings demonstrates that evaluations really matter.

Issue No. 3: **SETTING REALISTIC EXPECTATIONS**

Unrealistic expectations, for example in relation to timing and scope of evaluation findings, can cause disillusionment. This was the case with earlier evaluation efforts in many countries, for example in Sweden in the 1960s and 1970s. The introduction of programme budgeting and evaluation ran into difficulties due to unrealistic expectations and insufficient interest and support by ministries. Recognition of the inherent limitations of evaluation is important.

Politicians who put resources into evaluations do expect their findings to be used to improve decision-making in a practical and timely manner. Evaluators, on the other hand, may stress the longer-term value of cumulative findings that may take years to emerge. Policy-makers are therefore looking to evaluators to provide information that the evaluators say they cannot provide. Still, evaluators continue to promote more and more evaluations.

These differences between the demand and the supply sides can lead to distrust and disappointment unless both parties are able to clarify their expectations and establish what evaluation can realistically offer. Potential users need to be aware of the likely limitations of evaluation, and be able to create effective demand that can realistically be satisfied. The objective of evaluation is to provide useful information and reduce uncertainty in relation to the policies it studies, but it cannot give “right” answers or replace judgement in decision-making. Moreover, the type of information produced through evaluation will depend on the choice of evaluator (see later discussion). The commissioner of an evaluation has to act as an educated and critical purchaser and be able to avoid incompetent evaluators with overly optimistic views of the project.

Issue No. 4: **SYSTEMATISING EVALUATION ACTIVITIES**

It is often claimed that a certain level of institutionalisation has to occur before evaluations can play a genuine role in the management of government organisations. Institutionalisation aims at creating conditions for a high-quality supply of evaluations, the utilisation of evaluation findings, and organisational learning.

An important aspect is to integrate evaluation efforts strategically with the overall performance management framework. Various review and control activities have to be built into a logical, coherent framework to avoid duplication of activities and too burdensome reporting requirements. The degree of systematisation can vary considerably:

Degree of systematisation		
	<i>Informal</i>	<i>Systematic</i>
• Requirements for conducting evaluations	→ voluntary and flexible	→ mandatory and regular
• Responsibility for conducting evaluations	→ decentralised	→ centralised
• Position of the evaluator	→ internal	→ external
• Reporting	→ to the executive branch of government	→ to the legislative branch of government
• Focus of evaluations	→ improved decision-making and implementation	→ accountability
• Initiation	→ initiated and carried out locally	→ initiated and carried out centrally
• Openness	→ low	→ high

The way evaluations are organised affects their role in decision-making. Integration of evaluations with different policy- and decision-making processes is a key aspect of ensuring their use (see later discussion on linking with the budget process).

The value added and the costs of evaluation compared with, for example, ongoing performance measurement have to be assessed. Some governments have earmarked money for evaluation purposes (e.g. the evaluation of EU expenditure programmes is normally financed from the programme budget, a small part of which is often earmarked for this purpose) or created special funds for financing evaluation projects.

In Australia...
A tension between flexibility and control in systematising evaluation has been solved under current arrangements by institutionalising the <i>process</i> while largely allowing individual departments to decide <i>how</i> to do the job. Department of Finance can participate in selected major evaluations and help provide quality assurance.

Issue No. 5:
LINKING WITH THE BUDGET PROCESS

For evaluation to be effective it must be an integral part of policy design and implementation. Formal and systematic linkages with the budget process are therefore important. However, a comparison made of management practices in seven countries concluded that in none of them were budgeting, auditing and evaluation properly integrated (Gray & Jenkins, 1993).

Expectations in relation to the use of evaluation findings in the budget process should be realistic. Budgeting involves political choices, and evaluation is only one of many factors to consider. It is useful to develop a constructive relationship between the evaluator and the organisation being evaluated. Some consider that linking evaluation with resource allocation will reduce the trust between the parties, and consequently the quality of evaluation. Control- and savings-driven evaluation, where evaluation findings are only used as a basis for budgetary cuts, will create resistance. Involving the managers and staff of the organisation being evaluated in the process can improve the balance of findings and create trust.

A number of countries (for example Australia, Sweden) and also the European Commission have made attempts to improve links with the budget process. The examples presented below demonstrate some of the practical solutions.

In Australia...

In Australia, evaluation has been linked with the budget process. Programme budgeting, including programme evaluations, provides a framework for decision-making. Evaluation has been a principal means of reallocating new resources on the basis of programme effectiveness.

A review of the use of evaluation in the 1994-1995 budget process found that evaluation findings influenced 77 per cent of new policy proposals and 65 per cent of savings options. Evaluation was assessed to have influenced Cabinet deliberations in 68 per cent of the proposals considered. This was a significant increase over the results of the survey over the preceding budgets, partly reflecting the fact that one large package of proposals, the Working Nation package, which was heavily supported by evaluation results, comprised a large component of new policy proposals in the budget.

The Australian position is that evaluations should be conducted by those who manage the programmes, but that increasingly the overall process should be guided by an integrated approach to performance management based on good practice principles that is co-ordinated by those responsible for allocating resources. The Department of Finance has a central role in co-ordinating and guiding evaluations conducted in departments and agencies, and does sometimes participate in individual evaluations.

In Sweden...

At the end of the 1980s, a new budget process was launched in Sweden. One of the aims was to make previous *ad hoc* evaluation a formal part of decision-making processes. In addition to producing annual results information, agencies were required to carry out in-depth reviews of their activities every third year. The objective was to facilitate discussions on policy outcomes and provide a better basis for important policy decisions.

The evaluation system has recently been revised to allow for more flexibility. The main document used in budget process is the annual report containing mainly output information. In-depth reviews are only carried out if the respective ministry asks for them. The briefing material for the budget process has improved greatly. However, no systematic information on the use of evaluation is available. The general impression is that information presented in the agencies' annual reports is mainly used for control by ministries and for fine-tuning of policies.

In the European Commission...

A policy of systematic programme evaluation was adopted in 1996. Operational services are responsible for carrying out evaluations, whereas the Directorates General for budget and financial control support and control the quality of evaluations.

An important link between budgeting and evaluation is created through the requirement that all programme proposals must be accompanied by a financial statement which sets out an evaluation plan for the programme. Proposals for new programmes must be based on *ex ante* evaluation, and proposals for programme renewals are to be supported by evaluations. In short, there is "no appropriation without evaluation".

Issue No. 5:

CHOOSING THE APPROPRIATE EVALUATOR

The organisation of evaluation activities reflects the role governments want evaluation to have. Institutional arrangements affect what can be expected from evaluation; they influence the focus of evaluation, its links to decision-making and its overall status. This issue involves a number of aspects.

First, it is possible to give (one or more) government organisations a formal mandate to carry out or commission evaluations. Second, the government can promote development of evaluation capacity, both within and outside organisations with a formal evaluation mandate. The third aspect is the choice of an appropriate evaluator for each evaluation.

It is common to give government organisations a formal mandate to carry out or commission evaluations. It is common to give Audit Offices a mandate to carry out evaluations or performance audits. There are examples of requirements to set up evaluation units within ministries. This can be done across the board or in individual ministries. There are also examples of semi-independent evaluation units, both with a general and sectoral evaluation mandate.

Development of evaluation capacity does not have to be limited to those organisations that have a formal mandate to carry out evaluations. Many governments consider it important to having a range of different evaluators (both within and outside government) to choose from.

The choice of evaluator is an important factor in the effectiveness of evaluations. It is essential to analyse the objective of each evaluation. For example, internal evaluation will probably address relevant issues and come up with practical recommendations, but it is unrealistic to expect it to provide critical information and address issues in relation to whether a programme should be continued or not. Independent audit offices can address more fundamental issues in relation to outcomes, but their choice of issues, findings and recommendations are not necessarily politically relevant. Semi-independent evaluators may be able to combine these two perspectives.

Evaluations can be carried out by a range of actors, and only some of them will have a formalised mandate to carry out evaluations. Evaluations can be conducted internally in the administration, externally by independent organisations or, as is often the case, by a mixture of different actors. The degree to which these actors can be labelled as internal or external evaluators will depend on the perspective; all but self-evaluation is external to the programme evaluated, evaluation units in ministries are internal to the ministry but external to the programme, audit offices are typically fully independent, etc. The following table shows a range of organisations that government can potentially give formalised evaluation mandates or choose as evaluators. The table also gives an indication of which of these can be seen as internal, external or independent. Depending on the perspective there is an overlap between these categories.

Positions of Evaluators and Their Relation with the Evaluated Programme		
Relation with the evaluated programme	Evaluator	Position of the evaluator
Internal	Organisation evaluated	Inside the administration
	Special evaluation units within ministries	
	Central management agencies	
	Commissions	
External	Management consultants	Outside the Administration
	Academic or research institutions	
Independent	Community and consumer groups etc.	
	National audit offices	

When evaluation is *carried out by the organisations evaluated*, the assessment contributes to the fundamental information needs of the organisation. Internal evaluators know the programme and issues they study, have an easy access to the ongoing information produced in the organisation and also possess important informal information. Data can be relatively easy and inexpensive. Self-assessment enhances learning from experience and facilitates putting results to use. Also, requirements are often less demanding because complex issues in relation to the relevance and outcome of the programme are often not the subject of these evaluations.

The problems of in-house evaluation concern the time and skills of the staff, the rather limited range of issues that are addressed, the possible lack of objectivity, and consequently the lack of credibility of findings. This can result in limited use of evaluation outside the organisation. However, this is not necessarily a problem if the evaluation is aimed mainly at improving implementation, rather than evaluating the overall relevance of a programme.

In Canada...
Departments and agencies are responsible for evaluation of programmes. Evaluations are often carried out by specialised corporate evaluation units based in departments. They are sometimes criticised for not adequately reviewing the ongoing rationale of programmes. In 1996, a major programme review was undertaken, driven by ministers through a special committee established for this purpose. Such major evaluations require clear political leadership.

In Sweden...
Government agencies are expected to evaluate and report their performance and results. A study carried out by the Expert Group on Public Finance in 1996 found that the quality of public agencies' self-evaluations was generally fairly poor: conclusions drawn were often not supported by the evidence, etc. Quality control mechanisms, such as independent control groups, were called for.

Evaluation by *central agencies* is able to address government-wide issues and cross-sectoral programmes, study more profoundly the impact of different programmes, and possibly more rigorously review the ongoing relevance of programmes. Also, the authority of the evaluator and access to key decision-making processes may enhance the use of the evaluation. The perceived reliability of findings and the more political nature of the issues addressed often raise interest outside the evaluated organisation and facilitates the use of the evaluation in the political decision-making process.

Problems may arise due to distance from the programme being evaluated -- central agencies have less direct access to information, and data gathering may be limited by time, resources and other constraints. Evaluators from central agencies may have conflicting roles (e.g. providing advice to both individual ministries and agencies, as well as the Ministry of Finance). Credibility may be at risk, unless the evaluator is able to clarify its roles in relation to individual evaluations and make a distinction between objective information and judgements that may be influenced by the mandate of the organisation (for example value for money). In addition, programme managers may not be eager to accept findings from external evaluations and recommendations may not be easy to implement.

Programme managers, or central agencies, often act as the co-ordinators of evaluations carried out by external evaluators (for example research institutions and consultants). This co-ordination function involves selection of the evaluator, the contracting process, selection of programmes for evaluation, and assisting evaluators in conducting the evaluation.

In France...
The inter-ministerial committee on evaluation decides the number of evaluations conducted in response to requests from the Cabinet, agencies and the Audit Office. The independent scientific council advises the committee on methodology and quality of evaluations. Evaluations are funded through the national fund for the development of evaluation.

In the Netherlands...
Internal budget directorates and the Ministry of Finance are responsible for co-ordinating, promoting and determining the use of evaluation, and ensuring that necessary advice, guidance and research expertise is provided.

In Australia...
The Department of Finance has a central role in assisting ministries and agencies to become familiar with evaluation, to develop evaluation plans and activities, to develop expertise and reference material, to train staff, and to issue a register of completed evaluations. It also participates in some evaluations with the individual ministries.

When evaluation is *carried out by external evaluators*, their autonomy and evaluation skills often improve the range of issues addressed and the reliability of findings. This in turn widens the interested audience. On the other hand, external evaluators may have less understanding of the substance and the political and administrative culture of the evaluated organisation. In addition, some external evaluators may be willing to offer rather standard evaluations or present only findings that please the customer. Academic institutions may tend to produce overly theoretical studies.

Some suggest that for evaluation to succeed it has to be *anchored in several places* (Mayne, Divorski, Lemaire, 1998). Only then can it meet differing demands and provide information for a multiplicity of markets. There is a risk of unproductive competition and fragmentation of evaluations, but using different evaluation suppliers will broaden the issues and perspectives covered in evaluations and thus increase the possibilities for using their findings. By diversifying the supply, it is possible to combine the benefits and minimise the limitations of single evaluators. This can also be done by creating semi-independent evaluation bodies.

Because developing evaluation capacity in many different organisations is a very challenging task, it may be wise in the early stages to place evaluation activities where there has been some evidence of demand for them, and preferably where resources are available as well. Some suggest it may be worthwhile to centralise evaluation capacity to some extent in the beginning -- it can be expanded to other locations when skills and other resources increase (Boyle, 1997).

Several studies and experience suggest that government organisations appear more receptive to information produced internally, while external evaluation is often seen as threatening (e.g. Rist, 1994). However, the need to have evaluation accepted and hence implemented has to be balanced by proper quality control mechanisms to enhance the credibility of internal evaluations. This can be achieved by increasing their transparency and publicity, having external reference groups and conducting reviews (see later discussion on maintaining technical quality).

The extent to which “meta-evaluation,” or the evaluation of evaluations, occurs and where it is located is also relevant. For example, there may be advantages in promoting internal evaluation that will subsequently be reviewed by external bodies.

Issue No. 7:
PLANNING EVALUATIONS

The good preparation of evaluations is a major factor in their success. It is worth putting time and effort into selecting the topic, determining scope and objectives, planning the evaluation process and choosing criteria to be used. Careful planning makes the management of evaluations easier and contributes to the quality of their outcome. However, planning should not become too rigid; a degree of flexibility allows evaluations to be adjusted to changes in needs and circumstances.

The commissioner, often in consultation with evaluation users, should take the final responsibility for the choices in relation to evaluation planning. Attention has to be paid to defining objectives, issues, criteria, available information, methods of data collection and methods of analysis. The timetable, budget, evaluators and the overall organisation of the evaluation must also be decided on. The way in which the results are to be presented, used and monitored has to be considered.

Planning also involves identifying users of the evaluation findings and recommendations, and defining the responsibilities of different actors and how they are informed and motivated. The use of evaluation findings needs to be considered, including any concrete steps towards the implementation of recommendations (see later discussion on monitoring or follow-up).

Careful planning of evaluations strengthens the commitment to act upon the findings and may help to ensure that further action will be taken. Defining the users of evaluation in the planning phase greatly influences evaluation design. Information needs of users have to be addressed and the interest of those being evaluated considered in order to ensure their involvement and co-operation.

In addition to planning individual evaluations, it may be useful to compile and publish departmental or government-wide evaluation plans (i.e. the Portfolio Evaluation Plans in Australia; the Annual Evaluation Programme of the European Commission). These inform stakeholders of what subjects are to be studied and with what scope and objective, and contribute to the co-ordination and implementation of evaluation.

Issue No. 8:
TIMING EVALUATIONS APPROPRIATELY

To be of use, evaluation has to fit into policy and decision-making cycles. Timeliness is often a major problem in evaluation. Evaluations have to be planned to deliver information when decisions are being made. However, timing can be difficult because decision-making is not very predictable and policies are often revised before their outcomes can be assessed. Different users of the same evaluation often have separate timetables. Evaluation may involve surprises and it is better to be too early than too late with the results.

Timing is also important from the point of view of the evaluated programme. It is essential not to attempt to conduct *ex post* evaluation before the programme has achieved a certain degree of maturity. However, good evaluation is best supported by an efficient monitoring system which is put in place once the programme is started. It is also important during an evaluation to allow enough time for analysing the

material. The commissioner needs to monitor the evaluation timetable carefully so as to act in time if problems arise.

Issue No. 9:
MEETING USER NEEDS

Evaluations must be targeted at a certain audience and meet the information needs of users. The status, perceived reliability and connections of the user influence the credibility and hence the use of evaluation. Since evaluation is essentially oriented towards the policy- and decision-making processes, it is important to be familiar with the user's way of thinking and acting. Understanding the process of political interaction is as important as being able to undertake sound analysis. A dialogue between the cultures of users and conductors of evaluation is needed. The different decision-making processes and the variety of users of evaluations may also require the diversification of evaluation activities (e.g. in Sweden there are several suppliers of evaluation briefing different decision-makers).

However, evaluation may have roles that go beyond direct service to specific users. Evaluation must also recognise other responsibilities. Ethically unacceptable behaviour, misinterpretation or misuse of evaluation findings on the client's behalf must not be accepted.

Issue No. 10:
ENSURING RELEVANCE

It is important to assess the possible benefits of evaluating a particular programme and the degree to which reliable information can be gathered about its results. Evaluations should address questions or study programmes that are significant for budgetary, political, management or other strategic reasons. This is especially important in the early stages of evaluation within a government, as even a few high-profile evaluations may have a significant impact on the evaluation culture. For evaluation to be worthwhile, the programme assessed should also be reasonably stable. Any requests for evaluation have to be honest: there should at least be a possibility of evaluation findings being used.

It is often most difficult to evaluate issues that are the most interesting from a political perspective, for example questions of the outcome and overall policy relevance. Sometimes issues of minor importance are addressed because of insufficient interaction between the evaluator and users. Choosing the evaluator carefully and possibly creating a reference group will facilitate this interaction. The commissioner of evaluation should be sure to spend enough time defining the terms of reference for the evaluation and supporting wide discussions about the terms before the evaluation is started.

A degree of flexibility is necessary to maintain relevance throughout the evaluation process. It may be necessary to adjust an evaluation to changes in the evaluated programme or changes in the demand for information. The relevance of particular issues may change as the evaluation proceeds. Therefore, evaluation contracts should not be too rigid. The evaluator should be proactive and not confine the evaluation to the original issues if other perspectives turn out to be relevant. However, a certain balance is necessary as too flexible evaluations run the risk of being captured by the evaluator.

In France...

According to some commentaries, evaluation studies commissioned through the Scientific Council for Evaluation are mainly limited to topics of relatively minor importance. The reluctance of bureaucracy to address politically sensitive issues, and the avoidance of conflicts in some political circles, undermine the relevance of evaluation. (Boyle, 1997; Toulemonde, 1998). Also the cultural gap between the academics conducting evaluation and the politicians receiving the information has limited the use of evaluation findings.

Issue No. 11:
INVOLVING STAKEHOLDERS

Stakeholders are those who have an interest in evaluation because the process or findings may be relevant to their role or action. Stakeholders can be policy-makers (parliamentarians, ministers, other decision-makers), the budget department, managers of the programme, civil servants implementing the programme, the target group of the programme or other interest groups interested in or affected by the programme.

There are different views on the desirability of a participatory approach to evaluation. In any case, a strategy to manage the stakeholders is needed. Defining stakeholders and involving them in the evaluation process will usually increase their commitment and willingness to use evaluation findings. The different perspectives contributed by different stakeholders can enrich the evaluation process.

However, involving stakeholders can have drawbacks. It can be difficult to manage and it can be difficult to balance varying interests. Some stakeholders (including the managers and staff of the programme being evaluated) may be reluctant to participate and provide information, especially if they think that the evaluation findings may be critical of their performance. Managing relations with stakeholders can be more costly and time-consuming than a non-participatory approach. In addition, involvement by stakeholders may decrease the credibility of evaluation.

Despite these difficulties, a partnership among stakeholders is useful. Otherwise, evaluation risks becoming a bureaucratic, formal procedure carried out by professionals having only weak links to potential users.

Partnership can be established between all parties: commissioners, evaluators, evaluatees, potential users and other interest groups. Involving the stakeholders can also be made a formal part of evaluation by creating steering groups or advisory committees. The scope of co-operation may vary, for example according to the phase of the evaluation process. Partnership in defining the terms of reference is often considered most beneficial for maximising the use of evaluation. Some feel that stakeholders should also decide the methodology applied and be involved in analysis and interpretation of results.

Evaluators must pursue a fair and confidential interaction with the parties involved. They have to be aware of the balance of interests to be served during the evaluation process. One stakeholder group that often needs to be given greater attention is the users of the services or the target group of the evaluated programme. This naturally complicates the evaluation, since the outcomes of the programme have to be reviewed not only in relation to programme objectives but also in relation to the needs of customers. It is sometimes suggested that evaluators should also particularly safeguard the perspective of any minority groups involved.

In Sweden...
<p>The commission system has played a central role in preparing political decisions during this century. Although the relative importance of the system has diminished in recent years, almost 200 commissions have been working annually. Commissions combine the views of stakeholders by bringing together the representatives of political parties, the academic world, trade unions, employers' associations, consumer organisations, environmental interest groups, etc. In addition, reports produced by commissions are referred to a large number of interested parties -- both public and private -- for consideration and review. In this way the expertise available in society is combined in order to analyse and prepare political decisions.</p>

Issue No. 12:

ENSURING METHODOLOGICAL QUALITY

The methodological quality of evaluation encompasses issues such as relevant criteria, adequate evidence and hence reliable and clear findings. The soundness of methodology may have a major effect on the credibility and use of evaluation. This aspect requires special attention since the market for public sector evaluation is growing rapidly in many countries.

To some extent there may be a trade-off between methodological rigour and the utility of evaluation. Evaluations that are dominated by complicated research methods and do not make any conclusions unless they can be proved may not be the ones that receive the attention of policy-makers. Also, it can be very challenging to integrate use of scientific methods on the one hand and pragmatic solutions to policy concerns on the other. The issues of cost and time often limit the scope for using the best methods, as they may be too costly and time-consuming to be a realistic option. In short, pragmatic considerations have to be balanced with epistemological ones.

Methods should be selected, to the degree possible, in the planning phase of evaluations. This includes methods for gathering data, use of both qualitative and quantitative methods, sampling, causation, alternative explanations, and logic of argumentation. The conclusions of an evaluation should be based on the findings from the data, and recommendations should be based on the conclusions. Evaluators have to be explicit about assumptions, uncertainties and other methodological limitations of the evaluation. As far as possible, facts have to be separated from opinions or assumptions.

Quality control processes may vary. It goes without saying that excellent technical expertise and other resources contribute to a high-quality evaluation. The openness and transparency of the evaluation process allow external reviewers to scrutinise quality. Publication of evaluations can be a quality control mechanism, because more attention may be given to quality in evaluations that are to be published (see also later discussion on communicating results). Checklists can be created to set the quality standards for evaluations (for example the European Commission DG XIX Assessment checklist). Ethical standards for evaluation have also been adopted (for example by the American and Australasian Evaluation Societies) to enhance the quality of evaluations. It may be useful to create steering groups or totally independent external groups (for example the Scientific Council of Evaluation in France) to guide or review the quality of evaluations. Alternatively, the audit office may either have the responsibility or the authority to review evaluations if it wishes (for example the National Audit Office in Sweden). Reviews are especially important in enhancing the credibility and quality of internal evaluations.

Issue No. 13
MAKING JUDGEMENTS AND RECOMMENDATIONS

There are differing views about the value of making judgements and recommendations in evaluations. Some consider that judgements should be left to commissioners or users of evaluation. It is argued that the more general the conclusions of evaluation are, the easier it is to use them to shed light on different problems and thereby initiate discussion and a learning process.

On the other hand it might be suggested that judgements and recommendations attract attention, provoke and outline discussion and insist on decisions and actions to be taken. Decision-makers often have limited time and may be overloaded with information. Clear findings and recommendations, outlining main alternatives for action, are more likely to get their attention and provoke subsequent action. Also, constructive ways of suggesting improvements rather than focusing only on failures of a programme may limit resistance and contribute to a consensus on making necessary changes. In any case, it is necessary to make clear the criteria against which judgements and recommendations are made. Recommendations can be balanced by suggesting alternative solutions rather than focusing only on one. It can also be useful to present an implementation plan based on the recommendations in the evaluation.

Issue No. 14:
COMMUNICATING FINDINGS

Many evaluations do not have any effect because no attention is paid to communicating their findings. It is essential that evaluation findings are presented in an effective and timely fashion.

Information policies naturally vary, but the credibility and use of evaluation is often increased by openness and transparency. In Australia, the results of evaluation are expected to be made publicly available. In Sweden the public has traditionally had access to all government material including evaluation reports. Evaluation findings, especially critical or negative conclusions, are often politically sensitive. Unnecessary controversy may be avoided by giving those evaluated some preliminary information about the findings before their publication.

Generally, openness is useful. However, depending on the particular situation, transparency may have positive or negative effects on the use of evaluations. Publicity may promote the implementation of evaluation recommendations by creating external pressure, but it may also raise resistance among those criticised and reduce their willingness to co-operate and take action. This may be limited by avoiding an accusatory tone and promoting the culture of learning. Instead of concentrating on shortcomings, the evaluation should report what is being done to overcome the problems identified. However, there may in some cases be a point at which publicity and particularly media exposure becomes counter-productive.

As discussed earlier, the timing of evaluation reporting is important in terms of matching the timetable of decision-making. Often, it is useful to make some preliminary findings available. A decision on whether to present recommendations as well as conclusions has to be made. Experience suggests that evaluations with political implications are often better received just after elections, when a new government is more willing to accept criticism and recommendations.

Publishing a formal report is the most common way to present the findings of an evaluation. It is essential that the report be clear, comprehensive and widely available. The structure and argumentation of an evaluation report has to be intelligible. An executive summary is useful as it is unlikely that all

stakeholders will read the whole report. In some cases it may be worthwhile to publish shorter versions of the report for various stakeholder groups.

Delivering the report is usually not enough. Subsequent explanation and marketing is also needed to establish a dialogue on the results of the evaluation. Meetings with various stakeholders can be organised to present and interactively interpret the findings and promote implementation of recommendations. Special attention has to be given to key decision-makers. Seminars, workshops and discussion groups can be organised.

Issue No. 15:
MONITORING OR FOLLOWING-UP

Monitoring the impact of evaluation promotes its use, guides the implementation process, and can contribute to improving future evaluation practices. Follow-up is also important in identifying possible additional evaluation needs.

Follow-up can be facilitated by an institutional framework which reviews the use of all ongoing evaluations, or builds a requirement for subsequent follow-up into each evaluation. It is important to allow enough time for implementation before any conclusions are drawn on the use or impact of the evaluation.

Sanctions for not using the findings of an evaluation or not following its recommendations are typically limited to publicity and persuasion. However, mere public attention may create pressure to implement programme changes.

Implementation of evaluation findings and recommendations often requires support and consulting services. This can be facilitated by financial and technical assistance.

Issue No. 16:
RECOGNISING NEEDS OF STAFF FOR TRAINING AND SUPPORT

For evaluation to succeed, it is necessary to have trained evaluators, well-informed commissioners and enlightened users. In many countries, the educational background of civil servants does not provide specific evaluation skills. Therefore, training is important, especially when evaluation is being introduced and developed.

Methodological, political, managerial, and communication skills are needed. Commissioners and users of evaluation have to improve their substantive knowledge of evaluation to be able to articulate relevant questions and develop evaluation objectives, manage projects and select evaluators. They must also understand the limitations of evaluation as well as its potential to improve decision-making. In addition to technical evaluation skills, a culture of ongoing review of public activities is important. External evaluators need to learn to understand the managerial and political context within which evaluation takes place.

Training and support can be organised by...

- creating networks within the organisation, the whole government and between evaluators and auditors of different fields (peer group networks such as the evaluation network in the European Commission, or more formal associations such as the European Evaluation Society, the Australasian Evaluation Society, the American Evaluation Association, the British Evaluation Society, etc.);
- publishing handbooks, success stories, newsletters, journals, etc.;
- organising seminars, workshops, training courses and visits, etc.;
- setting up demonstration projects;
- setting up a help desk, or similar, in the administration.

4. CONCLUSIONS

Epistemological and pragmatic considerations have to be reconciled in evaluation. It is essential that there be a match between sound evaluation capacity and a willingness to utilise the findings. Only then is programme evaluation able to have an impact on decision-making (Mayne, Divorski, Lemaire, 1998; Boyle, 1997). These prerequisites are interdependent so that improving one will further the other. This requires political will and systematic work.

It is recognised that the use of evaluation information is applicable throughout the policy cycle, not just after the programme is implemented or when policy comes up for periodic review. The crucial question is how governments can organise their evaluation activities to take full advantage of what evaluation has to offer. Although organisational learning is usually not a systematic activity, evaluation can greatly contribute to it.

There is no need to set rigid frameworks for the use of evaluation. As objectives and methods of evaluations vary, so does the use their of findings. The major impact of evaluations on organisations has sometimes been described as a longer-term impact on organisational management -- although direct programme improvements should also be sought.

Evaluation may not be appropriate for all decisions. Sometimes evaluation is too costly or time-consuming to be useful. Sometimes no options for action are available or there is no political will to make changes. And sometimes the decision has already been made on other grounds. Evaluation can be misused. It may be conducted only to legitimise political decisions already made. Inconvenient or politically sensitive findings are sometimes suppressed or the results are used only selectively. Although the involvement of users is an important part of evaluation there is always a danger of evaluation being captured by the client.

Despite these limitations, the essential precondition for increasing use of evaluation findings is the commitment and perseverance of those who believe it can be used to improve policy-making. They need to continue producing good quality evaluation information, working with the users and adapting it to the ever changing environment of the public sector.

Examples of the use of evaluation findings

Some examples of actual use of evaluation and suggestions for effective evaluation strategies are presented below. The degree to which evaluation findings have actually been used can be interpreted in two ways: one can either feel disappointed that not all evaluations findings are used or be satisfied that a major part of evaluations do influence decision-making. The latter view is probably a more reasonable one.

Sweden
<p>Although comprehensive studies on the use of evaluation have not yet been carried out, some remarks can be made. 90 per cent of evaluations are used as briefing material and are mainly used for adjustment of policies rather than as a basis for fundamental reassessment of activities.</p> <p>There may be long-term learning effects, but intentions of introducing more fundamental issues into the budget process have not been realised. However, the quality of budget briefing material has improved. Annual reports by agencies contain mainly operational data and serve primarily internal management and support control by ministries. The Parliament remains the least informed party in evaluation. (Furubo & Sandahl, 1993)</p>

Canada
<p>The evaluation policy and infrastructure have been developed over two decades. There have been financial and other restraints, but continued use of evaluation is a major indication of its usefulness. A large number of the evaluations have led to reallocation of resources and at least some use has been made of the vast majority of evaluations. A combination of departmental evaluation plans and evaluation committees make it difficult for evaluation to be overlooked.</p> <p>Studies suggest that about 20 per cent of evaluations have led to significant reform or termination of programmes, 40 per cent to their modification, 30 per cent to better understanding about them and 10 per cent to their confirmation. A 1993 study by the Office of the Auditor General found that evaluation managers attached the greatest importance to helping management solve operational issues and improve programmes. Much lower priority was placed on challenging programmes to improve resource allocation or on evaluating large programmes to provide accountability to Parliament (Mayne, 1994).</p>

The Netherlands
<p>In a study carried out by the Court of Audit to explore the scope, practice and use of evaluation, over 80% of the officials interviewed stated that direct use had been made of evaluation findings. When some of the evaluations were scrutinised more closely, twenty out of forty-two studies were used explicitly, i.e. there was some written evidence of their use. Only a few evaluations suggested major changes. (Leeuw & Rozendal, 1994).</p>

Examples of suggestions for an effective evaluation strategy

Ireland

Ireland is about to start systematising its evaluation activities. According to Boyle (1997), lessons from other countries suggest that a range of evaluators is needed (semi-independent evaluation bodies in the administration may be a good model). Evaluation priorities have to be set: comprehensive reviews conducted every three years may be too challenging in the early stages. User perspective has to be involved when possible. Training and development support is needed for evaluators and users, both "hard" and "soft" skills are important. Stimulation of effective demand -- sticks, carrots and sermons, earmarking resources, specification of questions by those commissioning evaluations -- is necessary. Finally, linkages between budgeting, strategic management and evaluation have to be established.

United States

Richard C. Sonnichsen has designed a model for an effective internal evaluation office contributing to the use of evaluation information. The model is based on research conducted in five federal government internal evaluation offices and his own personal experience. Organisational conditions for an effective internal evaluation office are sufficient resources, independence, reporting to a senior organisation official, agreement on goals and policies between the office and the organisation, sufficient rank for the head of the evaluation office in the organisation, organisationally experienced evaluators and their acceptance in the organisation as credible, independent and objective.

Process factors include a defined role and published agenda for the office, self-initiation of evaluations, effective written reporting and dissemination of evaluation results in the organisation, issuing recommendations, applying a follow-up procedure and having evaluation staff play an advocacy role in publicising results and recommendations. All these provide a proper combination of organisational conditions, evaluation skills and managerial attitudes needed for evaluation to succeed. (Sonnichsen, 1994)

Australia

The Australian experience to date suggests some critical success factors for evaluation: several public sector reforms which emphasise bottom-line outcomes and results, formal Cabinet-endorsed evaluation requirements (moving over time to a good practice principles-based approach to performance management), support of several key ministers, clear mandate to the Department of Finance and its advocacy of evaluation, support of ministry heads, sufficient momentum and consensus on evaluation being a worthwhile activity, networks of evaluators, central evaluation units in ministries providing advice, planning, co-ordination and quality assurance, and sound performance information systems. (Mackay, 1996).

REFERENCES

- BOYLE, Richard (1997), *Evaluating Public Expenditure Programmes. A Role for Programme Review*, Discussion Paper Number 1, Committee for Public Management Research, Institute of Public Administration, Dublin.
- CHELIMSKY, Eleanor (1985), (ed.) *Program Evaluation: Patterns and Directions*, American Society for Public Administration, Washington D.C.
- Commonwealth of Australia (1994), *Doing Evaluations. A practical guide*, Department of Finance, Australian Government Publishing Service, Canberra.
- DENTE, Bruno (1997), *The Policy of Public Policy Evaluation*; a Paper presented in Colloque international GAPP -- CNRS-ENS, Cachan, France.
- DERLIEN, Hans-Ulrich (1990), "Genesis and Structure of Evaluation Efforts in Comparative Perspective", in Rist, Ray (ed.) *Program Evaluation and the Management of Government*, Transaction Publishers, New Brunswick, London.
- Dutch Ministry of Finance (1996), *Quality Criteria, Responsibilities and Dissemination of Information on Policy Evaluation in Central Government Departments in the Netherlands*, AVT95/FZ39626.
- European Commission DG XIX (1996), *Assessing the Quality of Evaluation Reports*, DG XIX, Assessment Checklist.
- European Commission (1996), *Evaluation, concrete steps towards best practice across the Commission*, Communication to the Commission by Mr Liikanen and Mrs Gradin in Agreement with the President, Internal Document, adopted by the Commission on 8 May 1996.
- European Commission (1997), *Evaluating EU Expenditure Programmes. A Guide. ex post and Intermediate Evaluation*, DG XIX.
- FURUBO, Jan-Erik and SANDAHL, Rolf (1993), *Does the Results Matter? The Role of Stakeholders, Expertise and Executors in the Production of Result-Based Data*, Swedish National Audit Bureau, The Financial Management Department.
- FURUBO, Jan-Erik (1994), "Learning from Evaluations: the Swedish Experience" in Leeuw Frans L., Rist Ray C. and Sonnichsen Richard C. (eds.) (1994), *Can Governments learn? Comparative Perspectives on Evaluation and Organisational Learning*, Transaction Publishers, New Brunswick, London.
- GRAY, Andrew, JENKINS, Bill & SEGSWORTH, Bob (1993), *Budgeting, Auditing and Evaluation, Functions and Integration in Seven Governments*, Transaction Publishers, New Brunswick, London.
- GRAY, Andrew (1997), *Evaluation, New Public Management and Democracy*, a Paper presented in the European Evaluation Society Conference, Stockholm.

- GRAY, Andrew & JENKINS, Bill (1993), "Horses to the Water: Budgeting, Auditing, and Evaluation in Seven Governments" in Gray, A. Jenkins B. and Segsworth B. in *Budgeting, Auditing, and Evaluation. Functions and Integration in Seven Countries*, Transaction Publishers, New Brunswick, London.
- HATRY, Harry P. (1980), "Pitfalls of Evaluation" in Giandomenico Majone and Edward S. Quade (eds.) in *Pitfalls of Analysis*, International Series on Applied Systems Analysis, 8, John Wiley & Sons. Chichester/New York/Brisbane/Toronto.
- HOLKERI, Katju and SUMMA, Hilka (1996), *Evaluation of Public Management Reforms in Finland: From ad hoc Studies to a Programmatic Approach*, OECD, PUMA/PAC(96)5.
- HOUSE, Ernest R. (1980), *Evaluating with validity*, Jossey Bass, Beverly Hills.
- LEEuw Frans L., RIST, Ray C. & SONNINCHSEN, Richard C. (eds.) (1994), *Can Governments learn? Comparative Perspectives on Evaluation and Organisational Learning*, Transaction Publishers, New Brunswick, London.
- LEEuw, Frans L. & ROZENDAL, PIET, J. (1994), "Policy Evaluation and the Netherlands's Government: Scope, Utilisation, and Organisational Learning" in LEEuw, Frans L., RIST, Ray C. & SONNINCHSEN, Richard C. (eds.) (1994), *Can Governments learn? Comparative Perspectives on Evaluation and Organisational Learning*, Transaction Publishers., New Brunswick, London.
- MACKAY, Keith (1996), *The Institutional Framework for Evaluation in the Australian Government*, Paper delivered to a World Bank seminar on Australia's program and policy evaluation framework on institutionalising the process, Washington DC, 30 April 1996.
- MAYNE, John, DIVORSKI, Stan & LEMAIRE, Donald (1998), "Locating Evaluation: Anchoring Evaluation in the Executive or the Legislature, or Both or Elsewhere?", in Boyle, R. & Lemaire, D. (eds.) (1998), *Building Evaluation Capacity: Lessons from Practice*, Transaction Publishers.
- MAYNE, John (1994), "Utilising Evaluation in Organisations: The Balancing Act" in Leeuw Frans L., Rist, Ray C. & Sonnichsen, Richard C. (eds.) (1994), *Can Governments learn? Comparative Perspectives on Evaluation and Organisational Learning*, Transaction Publishers, New Brunswick, London.
- OECD (1995), *Governance in Transition. Public Management Reforms in OECD Countries*, Conclusions of the Public Management Committee.
- OECD (1997), *In Search of Result --- Performance Management Practices*, Paris.
- OECD (1996), *Performance Auditing and the Modernisation of Government*, Paris.
- OECD (1996), *Performance Management in Government -- Contemporary Illustrations*, PUMA, Occasional Papers, N° 9,. Paris.
- POLLITT, Christopher (1996), *Approaches to Evaluation*, Speaking notes for the Presentation to Key Evaluation Officials Within the European Commission, 28th October 1996, Brunel University, Uxbridge.

- POLLITT, Christopher & SUMMA, Hilka (1997), "Performance Auditing: Traveller's Tales", in Chelmsky Eleanor, Shadish William R. (eds.), (1997), *Evaluation for the 21st Century. A Handbook*, Sage, London.
- RIST, Ray C (ed.) (1990), *Program Evaluation and the Management of Governmen*", Transaction Publishers, New Brunswick, London.
- RIST, Ray (1994), "The Preconditions for Learning: Lessons from the Public Sector", in Leeuw, Frans L., Rist, Ray C. & Sonnichsen, Richard C. (eds.) (1994), *Can Governments learn? Comparative perspectives on Evaluation and Organisational Learning*, Transaction Publishers, New Brunswick, London.
- SONNONCHSEN, Richard C. (1994), "Effective Internal Evaluation: An Approach to Organisational Learning", in Leeuw, Frans L., Rist, Ray C. & Sonnichsen, Richard C. (eds.) (1994), *Can Governments learn? Comparative perspectives on Evaluation and Organisational Learning*, Transaction Publishers, New Brunswick, London.
- TOULEMONDE, J. (1998), "Constraints, Incentives and Culture-Building as Instruments for the Development of Evaluation Eemand", in Boyle, R. & Lemaire, D (eds.) (1998), *Building Evaluation Capacity: Lessons from Practice*, Transaction Publishers.

APPENDIX

REFERENCE GROUP ON EVALUATION

Australia

Sylvie Trosa, Department of Finance

Canada

John Mayne, Office of the Auditor-General

Sweden

Rolf Sandahl, National Audit Office

United States

Harry P. Hatry, The Urban Institute

European Commission

Hilkka Summa