



---

COURSE MANUAL

# E7: Policy Analysis and Implementation

Module 8 - Policy Implementation, Analysis and Evaluation

[Add institute name here]  
[Add School/Department name here]

# Copyright

© Commonwealth of Learning 2012

All rights reserved. No part of this course may be reproduced in any form by any means without prior permission in writing from:

**Commonwealth of Learning**

1055 West Hastings Street

Suite 1200

Vancouver, BC V6E 2E9

**CANADA**

Email: [info@col.org](mailto:info@col.org)

[Add institute name here]

[Add School/Department name here]

[Add address line 1]

[Add address line 2]

[Add address line 3]

[Add country]

Fax: +[Add country code] [Add area code] [Add telephone #]

E-mail: [Add e-mail address]

Website: [www.\[Add website address\]](http://www.[Add website address])

# Acknowledgements

The Commonwealth of Learning (COL) wishes to thank those below for their contribution to the development of this course:

<b>Course author</b>	Dave Barrows Associate Director, MPA Program Schulich School of Business York University Toronto, Canada
<b>Assisted by</b>	Quarrat-ul-ann S. Siddiqi, MPA Toronto, Canada
<b>Course designer</b>	Dave Smulders, MA Vancouver, Canada
<b>Course reviewer and revision author</b>	Robert Dibie, PhD Dean and Professor of Public Affairs Editor, Journal of International Politics and Development School of Public and Environmental Affairs (SPEA) Indiana University Kokomo, USA
<b>Educational designer</b>	Symbiont Ltd. Otaki, New Zealand

COL would also like to thank the many other people who have contributed to the writing of this course.

# Contents

<b>Module 8</b>	<b>1</b>
<hr/>	
Introduction.....	1
Implementation strategies .....	1
Monitoring .....	5
Performance measurement.....	7
Programme evaluation .....	10
Change and continuity in policy execution.....	22
Module summary .....	25
Self-study questions .....	27
References and further reading.....	28

# Module 8

## Introduction

The purpose of Module Eight is to examine the implementation, monitoring and evaluation of policy. The following topics will be discussed:

- Implementation strategies
- Monitoring
- Performance measurement
- Programme evaluation
- Change and continuity in policy execution

By the end of this module you should be able to:



### Outcomes

- *describe* the different approaches to implementation of policy.
- *explain* how monitoring policy operates and why it is necessary.
- *outline* the components and related concepts involved in programme evaluation.
- *describe* the potential impact of programme evaluation on policy, programmes and organisations.
- *analyse* authentic case studies in light of the information presented in this module.

## Implementation strategies

The traditional view of the policy process focused on the role of political institutions, such as legislatures and cabinets, and the impact they had on policy-making. The assumption was that the most important aspect of the policy-making process was the development of the policy itself. The implementation of policy decisions was assumed to be automatic and hence of little interest to the political scientist. Implementation, as implied by political scientists, fell into the realm of administration.

Since the 1970s political scientists have come to accept implementation as part of the larger policy-making process. The reason for the new-found interest in implementation was because policy-making in itself, especially in the United States and United Kingdom in the 1960s and 1970s, proved to be ineffective in bringing about fundamental and lasting changes in social reform attempts. Dunsire (1978) argues that governments had become better at legislating than at actually causing changes. He refers to the difference between the two activities as the implementation gap.



In trying to understand why policies fail, Hogwood and Gunn (1984) differentiate between non-implementation and unsuccessful implementation. Non-implementation occurs when a policy has not been put into practice as intended, which may occur for a variety of reasons. Unsuccessful implementation happens when a policy is carried out in full, as intended, but fails to produce the desired or expected results. Some policies may fail for both reasons.

Policy-impact studies indicate three main reasons for policy failure: poor execution, poor policy and just plain bad luck – that is, when external circumstances beyond the control of any individual or group of individuals result in policy failure.

Hogwood and Gunn also put forward different approaches to understanding implementation and improving its effectiveness:

- structural approaches;
- procedural approaches;
- behavioural approaches; and
- political approaches.

Let's take a look at each of these approaches in a little more detail.

## Structural approaches

Different types of organisational structures are appropriate for different types of organisational tasks and environments. As such, policy design and organisational design should be considered together in policy as much as possible. Hogwood and Gunn make a distinction between two types of organisations in the context of control over change and, by extension, policy implementation.

The first type of organisation is one where changes are generated within the organisation and/or are largely within the control of the organisation with respect to direction and timing. Planning and implementation in this case are seen as technical or managerial problems. Essentially, planning *of* the change occurs.

The second type of organisation is such that change and planning occur as a result of external forces, such as the environment or other organisations, or when the process of change is complex, causing difficulties in prediction, control or containment of the process. Here, implementation requires the use of an adaptive approach, whereby the policy-making process is non-linear and iterative. Essentially, planning *for* the change occurs.

Organisational structures involving planning *of* changes can be bureaucratic, with sets of clearly defined tasks within a hierarchical structure. Organisational structures where planning *for* change is necessary often yield organic structures, usually with less-defined tasks and relationships, and less emphasis on hierarchy. Organic structures are

especially useful when the organisation operates within a fast-paced environment, since there is a greater ability to adapt and greater capacity to process information when compared to bureaucracies in which rules about channels of communication and chains of command must be followed.

Organic structures are most appropriate for the implementation of those policies or programmes that are not just one-offs, single or limited in scope, but rather those that are subject to changes in policies over time. However, governments may be less content with organic structures, since there are demands for organisational stability, consistency and accountability which allow for comparisons and evaluations of activities.

### Procedural and managerial approaches

The development of appropriate processes and procedures is another important factor that has an impact on the effectiveness of implementation. With respect to technical or managerial problems within a bureaucratic organisation, there is a focus on processes and procedures for scheduling, planning and control. There is an assumption that a high level of control can be exercised over the process. Implementation in this case would include the following steps: design of a programme outlining tasks and objectives, performance measures, costs and timeline; execution of the programme; and the creation and use of monitoring and evaluation techniques, as well as appropriate measures to be taken if the programme moves off target.

Two managerial techniques summarise implementation within such organisations: network planning and control and programme evaluation and review technique (PERT). NPC “provides a framework within which programmes can be planned and implementation controlled by identifying tasks to be accomplished, the relationships between these tasks, and the logical sequence in which they should be performed” (Hogwood & Gunn, 1984, p. 211). PERT is a sophisticated form of networking, which allows for “probability estimates for the duration of each task, calculation of the ‘critical path’ along which any slippage would delay the whole project, monitoring of any ‘slack’ elsewhere in the network, and reallocation of resources to permit activities on the critical path to be completed on time” (Hogwood & Gunn, 1984, p. 211). A common use of network analysis by governments is for tasks such as scheduling building contracts by local authorities.

Where there is a low level of control over implementation and the results generated, the policy process itself needs to be broader. An approach of iteration rather than linearity to the policy process is required. In such cases the process would include the following steps:

1. forecasting, which involves creating different possible futures by altering assumptions;
2. contingency planning;
3. carrying forward options for as long as possible to allow for more information about possible successes to be accumulated;



4. tentative objectives and programmes;
5. monitoring, feedback and adjustment to programme; and
6. fundamental reviews of assumptions, objectives and programmes.

## Behavioural approaches

This approach advocates that human behaviour and attitudes also play a significant role in the implementation of policy. This approach also recognises that usually there is a resistance to change. General sources of resistance include the fear of change and resentment toward changes sought.

Fear of change can exist for a number of reasons. Usually it is because change causes uncertainty and ambiguity at varying levels, which not all people can tolerate. Also, there may be more specific fears, such as insecurity about one's specific position and role, including career prospects, within the organisation as changes are implemented. People may be uncomfortable adapting to change, which often requires learning and training, assuming different responsibilities and being evaluated against different or higher standards.

Changes in organisational structure as a result of a policy decision may cause resentment because of the effect that the changes would have on the bureaucracy and autonomy afforded to individuals and groups within the larger organisation. Moreover, the targets of the change may feel that the whole process of change has been hurried, without adequate time necessary to consider and evaluate the change or to transition into the new structures.

In order to avoid or reduce resistance, full information should be provided about proposed or expected changes as early on as possible to all parties that will be affected. In addition, there should be considerable consultation with the affected parties and as much of their participation as possible in decision-making. Such efforts would enable decision-makers to gain the trust and support of affected parties and allow the parties to express their fears and concerns about the changes, thus helping overcome one of the most difficult aspects of implementation.

However, some behavioural scientists suggest that a top-down approach to decision-making, which assumes that management knows best, is more effective. This would remove a large amount of time and energy spent on consultations and trying to achieve consensus.

Management by objectives (MBO) is another approach to decision-making. MBO combines management techniques with behavioural analysis. MBO acts to integrate specific objectives and their implementation. MBO requires that there be:

1. a hierarchy of goals so that individual managers can determine how their objectives fit into the organisation;
2. an interactive process to arrive at goals and objectives; and

3. a system of performance appraisal to evaluate the progress of management.

The MBP approach has been found to be difficult to apply in the public sector, primarily because public servants usually do not have direct control over the resources necessary for implementation. Furthermore, political intervention is common, as is the inter-governmental nature of the implementation of most policies. Thus a clear hierarchy of goals does not exist, making the approach difficult to apply in the public sector.

### Political approaches

A political approach to implementation considers the patterns of power and influence within organisations. The main argument presented in this approach is that unless power structures surrounding a particular policy and its implementation are incorporated into the planning of the organisation and procedures, the policy is unlikely to succeed. The political approach is based on the belief that there is a need for some dominant group (that could be made up of multiple groups) to impose its will. This is important given that most policies are created at one level of government but are actually delivered at the local level of government. Similarly, local governments can depend on higher levels of government and other organisations to initiate and authorise policies. Thus, there exists interdependence among different organisations, although different organisations hold different levels of power among the larger group as well as have differing levels of access to resources.

## Monitoring

The time to consider policy and programme evaluation is at the option selection and programme design stages. In fact, the ability to evaluate a particular policy option is often considered a criterion for option selection. The programme design stages should not only ensure that proper evaluations are built into the programme itself, but should also, to some extent, address how information and evaluation results are used.

In this next section you will consider the factors and problems that relate to the monitoring and evaluation of the policy process. The following topics will be covered:

- the need for monitoring;
- defining and measuring criteria for success;
- the need for information;
- the separation of programme effects from other influences; and
- the costs of monitoring.

### Need for monitoring

For evaluations to yield meaningful results, activities involved in delivering the policy (its implementation) should be specified and outputs



of the policy/programme should be identified as much as possible. However, the evaluation should not be limited to pre-determined criteria, since evaluating a programme against original objectives while the programme itself has not been implemented as originally planned could be misleading. Instead, the matter of specifying activities and outputs should be subject to continuous monitoring.

Effective monitoring requires that the delivery of the initially envisioned programme be specified, as it is necessary to have specific standards to measure programme delivery against. Programme goals should be linked to programme objectives. The degree of acceptable variance from intended goals and objectives over time should also be specified.

Monitoring involves the collection of information to help determine how well programme goals are being achieved. This could require that specific information collection procedures be devised since general administrative data may be insufficient in determining the extent to which goals are being met. Surveys, programme participant/client interviews and observation can also be used to monitor a programme.

Monitoring not only involves information collection, but also decision-making about what actions should be taken if performance deviates from the plan. Systematic failure to meet programme goals could be indicative of the need to redefine goals, or they may simply indicate that the policy objectives may be impractical to carry out.

### **Need for information**

A potential problem inhibiting the ability to effectively monitor is that information necessary to determine the impact of a policy or programme may not exist or may not be available in a useable form. While data collected throughout programme delivery can tell us about the programme recipients, it does not provide information about the programme's target population as a whole. Information may be available at a level that is too broad, such as national statistics, or at a level that is too narrow, such as local statistics. Furthermore, recording and administrative requirements vary across jurisdictions, potentially making comparisons difficult or impossible.

### **Separation of programme effects from other influences**

Another potential problem in monitoring is the difficulty of attributing particular impacts to a particular programme. For example, suppose students have just started receiving a programme aimed at improving their grades in a given subject. At the same time, they get a new teacher for that subject. This complicates an evaluation since it is unclear whether improvements made were due to the new teacher or the programme.

### **Costs of monitoring**

Systematic evaluations can be very costly, especially if using scientific methods for analysis such as the experimental approach, which will be discussed later in the block. Money spent on monitoring and evaluation

tools is a direct diversion from funds available for delivering services. This is especially critical for smaller agencies that deliver programmes.

Such agencies are less likely to have built up a critical mass of administrative resources to manage additional external information reporting requirements, such as those imposed by programme funders, especially when the tracking of such information does not appear to directly benefit the agency itself. For example, agencies may be required to make considerable investments in computing and tracking systems in order to meet administrative requirements to have access to government funds.

## Performance measurement

Kernaghan and Siegel (1999) define performance measurement as the process used to determine how well a government provides a service. This implies that results will be measured against some pre-determined standard.

Although performance measurement has been around since the 1970s, it has recently become even more important as governments continue to rely more heavily on non-governmental organisations to deliver government services. In such cases, the government needs to ensure that the non-governmental agency provides value for the money it is paid to render services. Thus, government or other funders may impose reporting on various performance measurement criteria, or at least require the attempt to link funding of the programme to performance measurement targets.

### Legislating performance measurement

The linking of funding to performance measurement results is being developed in many jurisdictions in Canada and the United States. The State of Florida is at the forefront of this approach. The phasing in of performance-based budgeting (PB2), including strategic planning requirements, began in 1994 under Florida's Government Performance and Accountability Act (GPAA), which legislated that PB2 be immediately piloted and all state agencies participate by 2002. The statutory provisions establishing the framework for PB2 include appropriations Acts specifying output and outcome measure for each programme and performance standards for each output and outcome, as well as the requirement that Legislative Budget Requests include programme performance measurements and results.

The State of Florida also offers incentives for performance to legislative guidelines. These include increased budget, personnel, flexibility and retention of unencumbered appropriations as well as employee bonuses and other resource improvements. Adverse actions for non-compliance in Florida occur in the form of budget execution and management restrictions. Disincentives include mandatory quarterly reports on progress, quarterly appearances, elimination or restructuring of



programmes, restriction or reduction of positions and reduction of managerial strategies.

However, service-providing organisations often disagree with the value of performance indicators, as they may not be appropriate for the type of service provided, or because they are beyond the control of the service provider. For example, policy for a highway system may use performance indicators such as the volume of traffic and accident rates. However, accident rates include those accidents that are beyond the control of highway management, such as individual recklessness or unexpected and harsh weather conditions.

## Performance measurement in practice

Kernaghan and Siegel (1999) present a useful example of performance measurement in practice by reviewing the Canadian International Development Agency's (CIDA) results-based management (RBM) programme.

CIDA developed the performance measurement programme to evaluate projects as a direct result of ongoing criticism by the Auditor General of Canada about the organisation's lack of ability to identify and measure the value generated by the organisation's funding for various international development projects.

CIDA's results-based management programme identifies the links between inputs needed for the project and anticipated project impacts. Tables 1 and 2 on the next page provide a breakdown of the terminology used in CIDA's performance measurement programme, as well as an application of the performance measures to an international development project in Thailand.

Table 1: CIDA's results-based management definition of terms

<b>Inputs</b>	Direct or indirect human, organisational, or material contributions to implementation of the project/programme.
<b>Activities</b>	Management, co-ordination, technical assistance and training organised and carried out by the personnel assigned to a project/programme.
<b>Outputs</b>	<b>Short-term</b> developmental changes that logically result from the implementation of project/programme activities.
<b>Outcomes</b>	<b>Medium-term</b> development changes that logically result from the production of project/programme outputs. Results of this type involve changes in partner institutions and must be defined so as to materialise within the project/programme cycle.
<b>Impacts</b>	<b>Long-term</b> development changes that logically result from the production of project/programme outputs and outcomes. Expected results of this type involve changes in the living conditions of populations in developing countries. It is not normally possible to achieve results of this type within the project/programme cycle.

Source: (Canadian Development Agency, Autumn 1997, pp. 3-4).

Table 2: CIDA's results-based management Thailand example

Inputs	Activities	Outputs	Outcomes	Impacts
Time of Canadian and Thai educators	Develop ten new courses to train workers in:	Train 800 workers in these courses	Workforce skilled in latest production techniques	Assist in the economic development in the eastern seaboard of Thailand with proper attention to environmental concerns
Time of Canadian and Thai educational development specialists	• Hazards materials handling	Develop twenty highly qualified instructors to deliver courses	Educational institutions able to plan curricula and develop and deliver additional courses in the future	
Travel costs for mutual visits to Canadian and Thai educational institutions	• Environmental management • Total quality management	Develop five highly qualified course developers who can develop additional courses in the future		
Computer time to develop course material	Establish analytical chemistry laboratory	Develop Thai educational institutions as centres for environmental testing	Assist local industry in meeting ISO 14000 environmental protection standards	
Books and other course material	Train ten chemists and technicians in analytical chemistry			

Source: (Kernaghan & Siegel, 1999, p. 177)



## Programme evaluation

Programme evaluations are generally understood to be periodic, independent and objective assessments to determine the adequacy of a programme given its objectives, design and results (Kernaghan & Siegel, 1999, p. 180). This section will look at various aspects of programme evaluations including:

- who conducts evaluations;
- types of evaluations;
- evaluation models; and
- sources of data for evaluations.

### Who conducts evaluations

Evaluations of programmes can be commissioned and conducted by different actors in the policy community. Groups that are either internal or external to a given agency may conduct a given agency's programme evaluation. Depending on who carries out the evaluation study, there will be different implications for the technical effectiveness and utilisation of the study.

#### Internal (operating staff)

The main benefit of having evaluation studies conducted by internal operating staff is that they have a good knowledge of the programme and what is involved in delivering it. This includes an awareness of the shortcomings of the policy design stage. However, operating staff may lack the expertise to carry out effective evaluations. Another drawback of operating staff conducting an evaluation is that a number of different organisations may be involved in delivering a policy – the examination of only one organisation within the group is insufficient to provide a well-rounded evaluation. Finally, recommendations suggested by the operating staff may be biased, since the operating staff themselves would have to implement the changes. However, for the same reason just mentioned, any recommendations made by the operating staff in the evaluation would also be more likely to be implemented.

#### Internal (specialised evaluation staff)

A specialised unit within an organisation can also conduct an evaluation. Such a unit is likely to exist within the same organisation as that responsible for programme delivery, but is separate from it, being more concerned with evaluation and analysis instead. The specialised unit may be more objective, with less of a vested interest in the results. A drawback in the separation of the unit from the rest of the organisation is that there may be tension or unease from the operating staff. As the unit is part of the organisation, objectivity of such evaluations can be questioned, especially by users of the results outside the organisation.

### External (commissioned by delivery organisation)

Numerous external groups or organisations may carry out commissioned evaluations. These include academic groups, commercial companies, such as management consulting firms and non-profit organisations. Commissioned evaluators are ultimately responsible to the organisation that employs them.

Commissioned evaluations may yield more objective and credible results than internal studies, especially if academics or experts in evaluation are hired for the job. Two reasons why evaluations completed by commissioned groups may actually be less objective are: the commissioning organisations may require the use of strict guidelines to be used in the evaluation, which may pre-determine the results and essentially act to legitimise the organisation's viewpoint; and outside research firms may provide rosier results than warranted to build contacts and relationships so as to be awarded future contracts (Hogwood & Gunn, 1984, p. 236).

### External (commissioned by funding or legislative bodies)

Legislative or executive bodies that have been involved in a particular policy's development may also commission evaluations. One of their main concerns will be whether the policy has achieved success. This may be important to legislative or executive bodies for two reasons: to judge a programme and its value, and thus assist in judging the policy and its value; and to become aware of the status of programmes, especially if they are unsuccessful, before opposition and other political parties or individuals are able to expose the problem (Hogwood & Gunn, 1984, p. 237).

## Types of evaluations

### Needs assessment

Sometimes a request for an evaluation is neither for the purpose of summative evaluation (to report on a programme to make funding decisions) nor for the purpose of formative evaluations (to report on what specific activities can be done to improve a programme). Instead a needs assessment that focuses on the aim of the programme itself is needed. Needs assessments usually result in revised allocations to needs that are of higher priority and revisions to existing programmes.

A needs assessment requires that the evaluator determine the weaknesses or problems of a programme that can be eventually improved and determine what future conditions may arise that will require programme changes to occur. Herman, Morris and Fitz-Gibbon (1997) say that needs assessments are often used to make implicit goals public and to critique or evaluate existing goals.

Needs assessments are required, since need cannot be assumed. They must determine the real users of a programme – recipients or the target market – and whether those users are really the ones who should receive



those services. Sometimes, the actual user group may not be the one originally intended by the programme developer to receive it.

While needs assessments can be similar to formative evaluations, in that they often result in changes to programmes, they typically differ in terms of the size and scope of the study. Typically, formative evaluators will dig deeper into weaknesses and may work directly with staff to implement improvements. In needs assessment, the study of need is the final product.

### Formative evaluation

Formative evaluations include any number of aspects to run a programme smoothly. A formative evaluation may even include a needs assessment. Formative evaluations may come in many forms including: special surveys, interview studies, progress testing and management studies. The formative evaluation usually occurs in the programme planning and implementation phases, since any recommended improvements and changes can be incorporated with greater ease. Attention in the study is given to monitoring implementation and achievement of goals. They are time-consuming because they are so broad that they often involve improvements to many aspects of a programme.

Formative evaluations result in changes made to staff, activities, training, as well as organisation and other programme materials. Changes are often incorporated during the evaluation itself.

### Summative evaluation

Summative evaluations collect and present information for summary statements that are used to judge the programme and its value. Evaluations can be commissioned by a variety of players, such as government, programme policy-makers, those funding the programme, or stakeholders. In a summative evaluation, an evaluator should try to provide some basis for comparison, or a benchmark to make judgements about the programme.

A summative evaluation should not be done when a programme has not been fully implemented, developed or has just recently commenced. Clear, measurable goals with ongoing activities and organisations are more suited for summative evaluations.

Problems with summative evaluations include: the expectation that a summative evaluator act also as a formative evaluator; the conflict of interest when funders of evaluations commission the evaluator; and the importance of rapport, trust and relationships, all of which affect the quality and impact of data collected. Summative evaluations usually result in decisions about whether to continue or discontinue a programme and/or expand or reduce it. See Table 3 to compare and contrast the characteristics of formative and summative evaluations.

Table 3: The differences between formative and summative evaluations

	Formative evaluations	Summative evaluations
Purpose	<ul style="list-style-type: none"> <li>improve programme efficiency and effectiveness</li> </ul>	<ul style="list-style-type: none"> <li>answer funding/audience questions</li> </ul>
When	<ul style="list-style-type: none"> <li>usually in planning and implementing phases</li> <li>could occur at any time during the life of a project</li> </ul>	<ul style="list-style-type: none"> <li>in accountability phase</li> <li>requires programme history</li> </ul>
Focus of activity	<ul style="list-style-type: none"> <li>more creative about looking for and making assessments and offering solutions</li> </ul>	<ul style="list-style-type: none"> <li>documenting or assessing programme effects and determining their causes and generalisability</li> <li>more interested in programme outcomes</li> </ul>
People	<ul style="list-style-type: none"> <li>evaluators work closely with staff and develop working relationships</li> <li>often carried out by internal staff</li> </ul>	<ul style="list-style-type: none"> <li>evaluators work with staff to a much lesser extent</li> <li>often carried out by external consultants and therefore appear to be more objective (but not necessarily so)</li> </ul>
Results	<ul style="list-style-type: none"> <li>can result in revising the legislation if changes in goals or processes are not covered by existing legislation</li> </ul>	<ul style="list-style-type: none"> <li>can have implications for future programme designs and decisions</li> </ul>

### Implementation studies

These studies focus on the materials, activities staffing and administration components of a programme and how it operates. Users of this study are interested in issues such as how programme requirements have been interpreted by programme planners across sites or descriptions of staff roles in a programme. This information can be used formatively to improve the programme. Other times it can be used summatively, where programme sponsors attempt to make decisions about whether the activities occurring are contributing to achieving programme goals. Often this occurs when the programme represents new theories on how organisations should be run or when the question of “can the programme be implemented” is the key.

### Outcome studies

This type of study examines the extent to which a programme’s most important goals are being achieved. These goals may be affective, cognitive (such as skill acquisition), performance or behaviour-based. Goals are stated in terms of participant outcome. An evaluator should also consider the unanticipated or unstated outcomes that arise from the programme.



This type of evaluation provides management with knowledge about:

1. the extent to which the problems and needs that gave rise to the programme still exist,
2. ways to ameliorate adverse impacts and enhance desirable impacts, and
3. programme design adjustments that may be indicated for the future.

Results of the study are used to refocus activities and goals and make changes in areas of failure. In this respect it acts in a formative manner. The outcome study can also behave in a summative matter if the results lead to the termination of a programme.

Some programmes have strict timelines for existence and serve a very specific time and purpose. Outcome studies would be used here. Sunset clauses, for example, require that unless something significant happens, it will be understood that the programme objective has been achieved by the sunset clause date and the programme can be terminated. However, programmes or political forces may have a bias to keep a programme running and show that objectives have not been achieved and that it is still required. As such, it is possible that an evaluation report may be biased by the implications that the study will cause.

### Accountability studies

These studies occur when a programme has become established and a permanent budget and organisational structure exist. They question the overall effectiveness of a programme's impact. An accountability study is a summative report, which answers questions such as whether to expand or contract a programme. This study is generally done to satisfy programme policy-makers.

In accountability studies, the evaluator represents the sponsor and/or the broader community and should not interfere in the programme or make direct suggestions for risk of tainting objectivity. Rather, data should be collected and a summary report written. The summary report describes the programme and its achievements, as well as making recommendations for improving future efforts and addressing public policy. While objectivity is desired, it is almost impossible to achieve for a number of reasons:

1. evaluators are asked to function in both summative and formative capacities;
2. funders hire evaluators; funders take "hard-line" approaches to programme results; and
3. the process of evaluation is subjected to working relationships with data providers.

## Evaluation methods

One of the biggest aspects of a programme evaluation is the determination of the methodology to be used. Posavac and Carey (1992) propose that the methodology for an evaluation should be determined after first identifying and meeting with relevant parties, assessing the evaluability of a programme and examining relevant literature. After these prerequisite steps, an evaluator should have enough information to make some methodological decisions about how to determine programme benefits.

Posavac and Carey go on to say that a programme that can be “introduced to some segments of a target population on a staggered basis... is more desirable than introducing the programme to the entire population at the same time” (Posavac & Carey, 1992, pp. 33-36). This approach enables evaluators to make multiple comparisons of the same group and others over time, in order to determine the success of a programme. Fitz-Gibbon and Morris point out that “without any comparison group, it is hard to know how good the [programme] results are, whether results would have been as good with some other programme and even whether the programme had any effect” (Fitz-Gibbon, Taylor & Morris, 1987, p. 35).

Furthermore, studying samples rather than an entire target population (especially when the population is large) for programme evaluation can be argued for on the basis of usually tight constraints on both time and money. The correct use of sampling should provide findings that can be extrapolated to the entire population, but at a significantly lower cost.

### Experimental method

When using the experimental method to determine the impact of the programme, it is necessary to set up control and experimental groups for analysis. The measurement of key items across both groups provides data that can be analysed to determine whether statistically significant results exist that can then be attributed to the programme.

A *control group* is a group of people who are as similar as possible to those in an experimental group, but unlike the experimental group, they do not receive any aspect of the programme being evaluated. They are measured in the same way and at the same intervals as the experimental group.

Ideally, a control group should be identical in all aspects to its corresponding comparison group at the outset of evaluation or data collection. However, in reality, control groups will always be different from the comparison group, to some degree. Two kinds of control groups exist to deal with this: the *equivalent* or *true control group*, and the *non-equivalent* or *comparison control group*.

### True control group

A true control group is one where participants are randomly selected for participation. Achieving randomness in selection is critical because any



results that are achieved by a programme are not likely to have been caused by something other than the programme. Randomisation also acts as a natural equaliser between the control and the experimental groups. Furthermore, it reduces the effect of any specific factor that has the potential to affect the results of the study, such as neighbourhood-specific illness causing withdrawal from a programme. Finally, random assignment to groups is an important factor in statistical analysis, which requires random assignment in order for proper statistical application and interpretation, in order to ensure the credibility of findings.

While randomisation is most effective in equalising groups when numbers are large, it is less effective as numbers get smaller and for very small groups (less than 15 people) other special considerations should be made.

### Non-equivalent control group

A non-equivalent or comparison control group is a group intentionally selected (non-randomly) specifically because it is similar to the comparison or experimental group. Where random assignment is not possible, a group is selected for comparison that is as similar as possible to the experimental group that will receive the programme.

Fitz-Gibbon and Morris list three important points about non-equivalent control groups. First, if an experimental group was selected using a particular methodology, then the comparison control group should also be selected by using a procedure that is as similar to it as possible. Secondly, both equivalent and non-equivalent control groups should be given the same major tests that the experimental group is given, since testing itself can make a difference to the evaluation by causing behaviour that otherwise may not have occurred. Examples include participants focusing attention on aspects perceived to be important in testing, getting practice doing tests, changing motivation levels due to testing performance and providing suggestions to improve the evaluation. Finally, differences and similarities between control and experimental groups should be reported and examined more in depth where necessary, to protect the integrity of the evaluation.

### Pre-post studies

Research designs outline what is to be measured and when. Pre-post studies are essentially before and after studies. Pre-tests are pre-programme tests given before a programme or an experiment starts. Similarly, post-tests are post-programme tests administered at predetermined intervals of the programme or at programme completion. Ideally, pre-tests and post-tests will be as similar as possible to allow for direct comparison of criteria measured.

Pre-tests can be used to:

- Select people for participation in study.
- Check assumptions that have been made in the planning process of a programme.

- Check to ensure the groups are comparable.
- Provide the grounds for checking what advancements have been made during a programme.
- Determine a more refined test for determining programme effects.

According to Fitz-Gibbon and Morris, pre-tests can be classified as attitude tests, achievement tests or ability tests. A pre-test as an attitude test occurs most often when a programme seeks to make changes to attitudes. Since a pre-test itself may influence responses to the programme and/or a post-test, it is recommended that in some cases only a randomly selected section of the group be given the pre-test.

A pre-test as an achievement test – especially when it is the same as or similar to the post-test – provides information on current status. In this case, the pre-test also serves to ensure that the control and experimental group are equivalent, which is particularly important when using a non-equivalent control group, a group of less than 15, and when there is a large variability in ability with the groups.

A pre-test as an ability test is often used when the post-test is to provide a measure of achievement. This way, post-test results can be related to ability. The effect of the programme can then be analysed with respect to differing levels of ability. However, a pre-test is more powerful when it is equivalent to the post-test.

Ability pre-tests are also useful when: knowledge about ability level will help to interpret the size of change that has occurred; a pre-test is not possible usually due to new information provided by the programme; the post-test will be an attitude measure and it is believed that attitudes will differ among different ability levels; or the programme is already underway and no pre-test has been given to either the experimental or control group and thus it functions as a retrospective pre-test.

Pre-tests should not be used when:

- Taking a pre-test would likely alter the participants in some immeasurable way (for example, when measuring attitudes).
- Using a pre-test would be meaningless (such as in programmes involving teaching a new language).
- A programme is already in progress and no pre-test was given before.
- The cost is too great in terms of time and money.

Tables 4 and 5 provide a graphical representation of the process flow for pre-test and post-test studies for true control group research designs and non-equivalent control group research designs.



Table 4: True control group research design

	TIME			
	Step 1	Step 2 (pre-test)	Step 3	Step 4 (post-test)
Experimental group	R	O	X	O
Control group	R	O	(C)	O

Source: Modified version of diagrams in Fitz-Gibbon & Morris (1987, p. 65).

Table 5: Non-equivalent control group research design

	TIME		
	Step 1 (pre-test)	Step 2	Step 3 (post-test)
Experimental group	O	X	O
Control group	O		O

Source: Modified version of diagrams in Fitz-Gibbon & Morris (1987, p. 65).

#### Where:

- R** = indicates Random assignment.
- O** = indicates a measurement of some kind, an Observation.
- X** = indicates the programme to be evaluated, the eXperimental programme, or the treatment given to an experimental group.
- C** = indicates the alternative programme given to the Control group, if any.

### Modelling

According to Freeman and Sherwood, without a “well-specified impact model”, the ability to control a programme’s quality and effectiveness is severely limited. An explicit impact model aids in “understanding how and why it [a programme] worked or for reproducing its effects on a broader scale in other sites and with other targets”.

An impact model, also called an intervention model, tries to develop hypotheses, often using existing theories or studies on social behaviour for the purpose of planning and implementing an effective programme within a measurable framework. However, Rossi and Freeman note that it is not uncommon to find that impact models are often reduced to “nothing more than assumptions underlying a programme’s operation... with little or no empirical basis... drawn from untested ways” conventionally used by practitioners. While these so-called “impact models” may result in some movement toward outcome achievement, the waste of resources can occur, as there will certainly be difficulty in the

measurement of efficiency and effectiveness of the programme, especially as it relates to the broadening or replicating of a programme.

Essentially, an impact model is a set of statements (guiding hypotheses) about the relationship expected to exist between a particular programme and its goal. It sets the method for reducing a gap between a goal and the existing condition – that is, the achievement of or movement towards a desired result. An impact model must have the following:

- the causal hypothesis;
- the intervention hypothesis; and
- the action hypothesis.

### Causal hypothesis

The correct understanding and determination of the causal hypothesis is crucial to the success of an impact model. It is the hypothesis that seeks to identify the influence of “one or more processes or determinants on the behaviour or condition that the programme seeks to modify. In order for the causal hypothesis to be useful, it must be “stated in a way that permits testing, or measurement,” referred to as the “operationalisation” of a hypothesis. The statement of causal variables in operational terms is an important part of an impact model since it identifies and brings the inter-relationships of influences and conditions to the forefront, thereby initiating the development of a programme framework in a manner that can be measured and evaluated.

### Intervention hypothesis

The intervention hypothesis explicitly states the relationship between the programme and what is going to be done. This hypothesis depends on the context of the specific “process or determinant” already identified in the causal hypothesis – that is, the existing condition that needs to be addressed. In other words, the intervention hypothesis indicates how a programme impacts on the cause(s) of a particular behaviour or condition that is expected to bring about a desired outcome.

### Action hypothesis

Another essential component of an impact model is the action hypothesis, which is necessary if one is to assess whether the intervention is necessarily linked to the outcome. An action hypothesis is required since the introduction of an intervention to simulate a naturally occurring chain of events may not result in the same behaviour and social processes that would have occurred naturally.

For example, Festlinger’s research into the attempt to reduce racial discrimination in the workplace by management training (knowledge being identified as a causal link) clearly illustrates the importance of the action hypothesis and the need for empirical study in this regard. The research found that those managers who exhibited the greatest change in attitude discriminated more than they had before – an exact opposite of what was expected. Empirical study and the development of a sound



action hypothesis can help reduce the chances of such a programme failure from occurring.

An impact model must also consider the issues of possible manipulation and feasibility. A programme intervention must focus on relevant variables that are subject to manipulation in practice, taking into account an identifiable time frame during which desired results can be achieved. Feasibility refers to the importance of avoiding interventions with low feasibility – that is, those programmes, which for one reason or another do not have social, financial or moral support. The existing social and political climate and their constraints must be appreciated in order to determine programme feasibility.

Other issues an impact model should take into account are the selection of a target market and the distinction between the group that will be subject to the programme and the total population that eventually will end up benefitting. The target population and what are believed to be its characteristics, focus the programme. As such, incorrect assumptions about the target population can require the need to make a major shift in a programme.

Ultimately, an impact model cannot be effective unless a well-developed implementation or delivery system is devised. However, a weak impact model is more likely to result in the implementation of a weak programme, thus increasing the risk of programme failure. Unfortunately, the implementation phase often takes the back seat to the design phase and reduces the potential to effectively achieve programme goals.

## Sources of data in evaluation research

According to Posavac and Carey (1992, pp. 45-50), there are five possible sources of data for evaluations. Let's take a look at each of the following sources in greater detail, considering some of their advantages and disadvantages for evaluation purposes:

1. programme records;
2. programme participants;
3. programme staff;
4. evaluator observations; and
5. community indexes.

### Programme records

Programme records and files usually provide an evaluator with reliable and inexpensive data. Archival data are most useful when they include objective data about the types of services offered, staff, workloads and trends at intervals over an extended period. In contrast, diagnoses and other subjective data may not be as useful or valid. Nonetheless, there are two main advantages of archival data: the participant is not affected by evaluator interference in measurement; and participants are captive in the records, meaning their data can be used without dealing (they can't refuse to participate).

In some cases, past performance reviews can be used to serve as control groups. Public documents are easily accessible and should also be examined. However, depending on the evaluation, health, education or social services records may also be reviewed, but care must be taken to maintain confidentiality. The evaluator must take steps to protect the privacy of individuals. Methods include coding data so that a person's name remains hidden and keeping master code lists in locations other than where the code is being used.

### Programme participants

Programme participants provide another, inexpensive yet important source of data for the evaluator. A person who receives a programme is well-positioned to assess many aspects of the programme. Recipients have direct contact with staff and knowledge of the programme that is unique. They will usually be the most knowledgeable about their current status with respect to the programme, and are the most in tune with their true feelings about the programme and its staff. Furthermore, participant self-assessments have been found to be just as accurate as other approaches for collecting behavioural and emotional data.

In some cases, however, participants may be incapacitated and unable to be useful sources. In other instances, participants may be able to provide good data for some programme aspects but not for others. The evaluator must motivate participants to spend time with them and to provide private information, along with their attitudes and judgments. Where possible, participant data or some proxy of it should not be overlooked. This type of information must be included in an evaluation if it is to be effective.

### Programme staff

Programme staff members are another important source of data. They include front-line workers trained to assess the condition of participants and administrators who have the inside scoop on how well the programme is managed in general and on a day-to-day basis. However, staff can be biased toward improvement and may have difficulty accepting failure.

Since a programme evaluation implicitly includes evaluation of staff to some extent, staff may have negative views towards evaluation and may be overly concerned about how the information they provide will be used. Hence they may not be forthcoming with information and support for the evaluator.

### Evaluator observations

Direct observations provide evaluators with another major source of data. The main advantage of evaluator observation is that evaluators are less-biased than programme experts, hence are able to be more objective in data collection and interpretation.



## Community indexes

Finally, community indexes are useful particularly in programmes geared toward improving community-level variables. Examples include a crime-reporting programme aimed at increasing citizen participation and reducing community-level indexes such as the arrest rate and crime rate.

The problem with community indexes is that they are subject to many unrelated variable influences that are beyond the control of staff. As such, an evaluator cannot determine success or failure of a programme using community indexes alone.

## Change and continuity in policy execution

### Use of evaluation results for decision-making

Hogwood and Gunn argue that, most often, evaluations do not result in a great deal of activity to occur on the part of policy decision-makers. Reasons for this may be that the evaluation study does not explicitly link policy implications with the research done, or that the policy implications are not obvious, making it difficult for the evaluator to address this issue.

The results may also be ambiguous, making it difficult to apply them. Furthermore, if the study is presented in overly technical language, decision-makers may not fully understand the study. It is advised that the research be communicated in a method that can be utilised by decision-makers.

Other reasons that evaluation results are seen to be essentially political and not utilised include: the results do not fit into the timetable for decision-making; there is organisational resistance to change to the implications of study findings; and the results, if critical, are difficult to accept since findings may cast doubt on the original judgement of programme sponsors, community members and delivery organisation.

**Note:** Evaluations do not determine the implications for a given policy. Political judgements about tolerable levels of variance from stated objectives are necessary.

### Programme succession and termination

Hogwood and Peters (1983) suggest that the policy succession will increasingly become a common feature of the policy formulation process in Western systems due to:

1. expanded activities by governments in the field of policy, such that most new policies will overlap with existing ones;
2. the need to change existing policies due to inadequacies or harmful side effects that become known; and
3. the limitations on the ability to scrap existing policies and start anew due to financial implications of existing policy commitments that may span a number of years.

If succession or termination does not take place, the policy is maintained and the policy continues. This can occur due to inertia, deliberate decision or failure of other meaningful options to be presented. The difficulties and costs associated with succession and termination may contribute to maintaining the status quo.

### Types of succession and termination

Let's take a look at four different types of termination and succession:

1. functional;
2. organisational;
3. policy; and
4. programme.

#### Functional

Functions are services that transcend individual organisations or policies, such as education or health. A number of different organisations can serve the same policy function. Functional succession in developed countries is rare, since this would require huge political effort. However, refinement of a function, such as eligibility for a social service support, may occur.

#### Organisational

Organisational succession or termination is easier to achieve, but organisations tend to take on a life of their own and seek survival regardless of the original purposes they were built to serve. More often than termination, organisations are transformed, merged or split up. Organisations are likely to survive programme cuts or policy changes.

#### Policy termination

The termination or succession of a policy is easier than terminating the organisations that carry it out. Hogwood and Gunn suggest that this is the case for a number of reasons. First, organisations, their management and staff have self-interest in the organisation's survival so that they can have jobs, status and power. Secondly, policies have fewer political allies and more critics than organisations. Finally, organisations may serve multiple policy objectives in addition to the policy considered for termination.

#### Programme termination and succession

Programmes refer to the specific measures used to carry out policy, such as training programmes that assist unemployed workers in finding work. Programmes are the most vulnerable to being terminated since they represent a smaller unit of analysis than functions, organisations and policies. At the same time, programmes are the closest to the problem and those affected, and their impact is easier to measure directly. Should programme results be poor, termination provides a relatively easy way to deal with criticism.



Policy succession can also take many forms, depending on the extent to which it is a replacement for an existing policy. In some cases, a successive policy may not cover the clientele to be served to the same extent as the old policy. Similarly, even if the policy remains the same, there could be a difference in coverage due to variations made to policy objectives. Policy succession can result in the same number of programmes or any number of different programmes.

Different types of policy successions will have different problems associated with them. Such effects should be included in the consideration of making policy changes. For example, when policy succession involves a reduction in the number of programmes, the process is likely to cause conflict among policy actors, since some organisations or groups will gain and some will lose. When two or more programmes or organisations are consolidated, intra-organisational problems will arise, whereas splitting up organisations will result in inter-organisational issues.

Finally, the policy analysis process would be incomplete without due consideration of policy and programme succession and termination. The problems facing governments are changing, resulting in a corresponding change to government activities. As a result, a cascade of changes in policies and programmes is bound to occur. Secondly, nothing lasts forever; changes in the prevailing wisdom about problems and solutions occur over time and such insights, knowledge and technologies need to be considered when reviewing or developing solutions. Finally, other significant changes, such as the amount and availability of resources, as well as the varying demands for service offered by the government, are bound to have considerable impact on policies and therefore deserve attention in policy analysis.

---

## Module summary



### Summary

The purpose of Module Eight was to examine the implementation, monitoring and evaluation of policy.

In the first section of the module you looked at four implementation strategies that can be used. These are: structural, procedural and managerial approaches, behavioural and political approaches.

The next section in the module covered the topic of monitoring within the context of the policy analysis and policy processes. Included in the discussion were: the need for monitoring, the need for information, the separation of programme effects from other influences and a consideration of the impact of costs on monitoring.

The concept of performance measurement was then outlined through the examination of an example of where performance measurement was legislated and an example of performance measurement in practice at a programme level.

The section on programme evaluation covered a number of sub-topics. First, four different groups that may conduct evaluations were described:

1. internal operating staff;
2. internal specialised evaluation staff;
3. external groups commissioned by delivery organisation; and
4. external groups commissioned by funding or legislative bodies.

Next the section looked at six different types of evaluations. These were:

1. needs assessments;
2. formative evaluations;
3. summative evaluations;
4. implementation studies;
5. outcome studies; and
6. accountability studies.

The section then went on to discuss three different evaluation methods:

1. the experimental method;
2. pre-post studies; and
3. the use of models.



Lastly, the section dealt with five different sources of data in evaluation research which were:

1. programme records;
2. programme participants;
3. programme staff;
4. evaluator observations; and
5. community indexes, in policy evaluation research.

Finally, the last section of the module dealt with change and continuity in policy execution. The section focused on the use of evaluation results for policy decision-making. More specifically, the section addressed programme succession and termination, as well as four different types of succession and termination methods. These included actions taken at the following levels:

1. functional;
2. organisational;
3. policy; and
4. programme.

---

## Self-study questions



### Study skills

1. Discuss the different approaches to implementing public policy? Why is it necessary to monitor a policy? Explain why some states have legislated performance measurement?
2. What are the most important factors to consider in conducting a political feasibility analysis? Take a specific example, such as raising fuel taxes to reduce reliance on imported oil, cutting mandatory prison sentences to lower the cost of keeping non-violent offenders in prison, or capping student tuition payments to permit greater access to higher education.
3. Discuss the different types of evaluation methods. Who are the major actors in the policy community that may be responsible for programme evaluation?
4. What are the main sources of data in evaluation research? Discuss the advantages and disadvantages of each of these sources for evaluation purpose.
5. Which of the many evaluation criteria are most important? Economic costs or efficiency? Policy effectiveness? Equity? Why do you think so? Are some criteria more important for certain kinds of policy questions than others?



## References and further reading



### References

- Canadian Development Agency. (1997, Autumn). *Results-based management: A guide for UPCD partners*.
- Doern, B. G., & Phidd, R. W. (1983). *Canadian public policy: Ideas, structure, process*. Toronto: Methuen.
- Dunsire, A. (1978). *Implementation in a bureaucracy*. Oxford: Martin Roberson.
- Fitz-Gibbon, C. T., & Lyons Morris, L. (1987). *How to design a program evaluation*. Beverly Hills: SAGE Publications.
- Herman, J. L., Lyons Morris, L. & Fitz-Gibbon, C. T. (1987). *Evaluator's handbook*. Beverly Hills: SAGE Publications.
- Hogwood, B. W., & Gunn, L. A. (1984). *Policy analysis*. New York: Oxford University Press.
- Kernaghan, K., & Langford, J. (1990). *The responsible public servant*. Toronto: Institute of Public Administration of Canada, and Halifax: Institute for Research on Public Policy.
- Kernaghan, K., & Siegel, D. (1999). *Public administration in Canada*. (4<sup>th</sup> ed.). Scarborough, Ontario: International Thomson Publishing.
- Norton, G., & Veneman, A. (2003, March 12). More than one way to protect wetlands. *The Washington Post*.
- Posavac, E. J., & Carey, R. G. (1992). *Program evaluation: Methods and case studies*. (4<sup>th</sup> ed.). Englewood Cliffs, New Jersey: Prentice Hall.
- Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach*. (3<sup>rd</sup> ed.). Beverly Hills: SAGE Publications.
- Sabatier, P. A. (1993). Policy change over a decade or more. In P. A. Sabatier and H. C. Jenkins-Smith, (Eds.), *Policy change and learning: An advocacy coalition approach*. Boulder: Westview Press.
- Sabatier, P. A. (1999). The need for better theories. In P. A. Sabatier (Ed.) *Theories of the policy process* (pp. 3-17). Boulder, Colorado: Westview Press.
- Taylor, D. W. (1991). *Business and government relations: Partners in the 1990s*. Toronto: Gage Educational Publishing.



Taylor, D. W., Warrack, A. A., & Baetz, M. C. (1999). *Business and government in Canada: Partners for the future*. Toronto: Prentice-Hall Canada.