

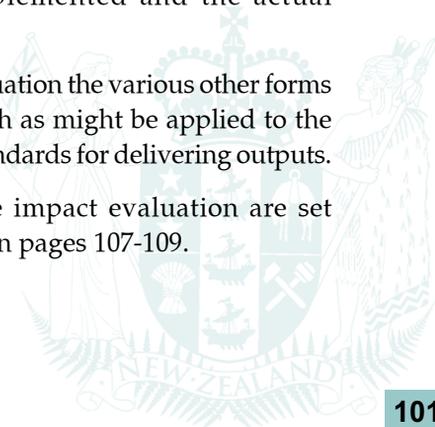
Introduction

Purpose of This Article

- 6.001 This article is intended to demonstrate the value of impact evaluation as a practical tool to enhance the quality of decision-making by the Government and Parliament. For impact evaluation to be valuable, decision makers need to:
- have a commitment to basing decisions on the best available information, consistent with their “world view” and the political context; and
 - accept that enhanced decision-making will contribute to optimising the effectiveness of government expenditure.

The Nature of Impact Evaluation

- 6.002 We understand “impact evaluation” to be a short-hand term for a particular form of performance assessment, the purpose of which is to:
- determine the actual outcomes from putting a policy into effect;
 - compare those outcomes with the desired outcomes when the policy was formulated; and
 - confirm or establish the causal link between the means by which the policy was implemented and the actual outcomes.
- 6.003 We do not regard as impact evaluation the various other forms of performance assessment such as might be applied to the achievement of performance standards for delivering outputs.
- 6.004 Our expectations for effective impact evaluation are set out in paragraphs 6.024-6.026 on pages 107-109.



A Survey and Case Studies

- 6.005 We conducted a survey to establish the extent to which government departments were using impact evaluation. We discuss the survey and its results in paragraphs 6.027-6.068 on pages 110-117.
- 6.006 To illustrate the subject, we present and comment on two case studies in paragraphs 6.069-6.130 on pages 118-130.

Other Commentary

- 6.007 We discuss the place of impact evaluation in the public management system in Appendix A on pages 131-136, and describe the current legislative framework in Appendix B on pages 137-139.

Our Objectives

- 6.008 Our overall objectives are to:
- develop the discussion of the subject that we began in our *Third Report for 1999 – The Accountability of Executive Government to Parliament*;
 - create a greater awareness among decision-makers of the practical value of impact evaluation; and
 - raise the expectations of decision makers and legislators that policy advice provided to them is informed by systematic, reliable, and relevant evaluation findings.
- 6.009 We do not expect that all Government policies will be evaluated. The overall net-benefit of such an approach would (in our view) undoubtedly be negative. However, it is desirable that, at a Government-wide level, a strategic selection of policy spending areas would be made (on the basis of stated criteria) to provide the basis for particular impact evaluations to be carried out. The resultant programme or schedule may include indicative timing for the evaluations to be undertaken over the medium to longer term.

The Significance of Impact Evaluation

6.010 In Part Four of our *Third Report for 1999 – The Accountability of Executive Government to Parliament*, we discussed issues affecting accountability for the outcomes of Government expenditure.¹ We framed the discussion in the context of the two critical questions that Parliament is presented with when determining whether or not to approve the Government’s expenditure proposals:

- What outcomes are the expenditure proposals intended to achieve (and should Parliament agree with them)?
- Is it likely that the proposed expenditure will achieve these outcomes?

6.011 In fact, the Government itself needs to be able to answer both questions in order to persuade Parliament of the soundness of its proposals.

6.012 Underlying those questions – and the ability to answer them – is the (apparently) simple concept of “cause and effect”, which in turn can be expressed in the form of three other questions:

- What policy objectives do we want to achieve?
- How do we go about achieving those objectives?
- Did we achieve the results we wanted to achieve, and did any unexpected results occur?

6.013 We can represent those three questions in the form of a “policy performance model” as shown in Figure 6.1 on page 104.

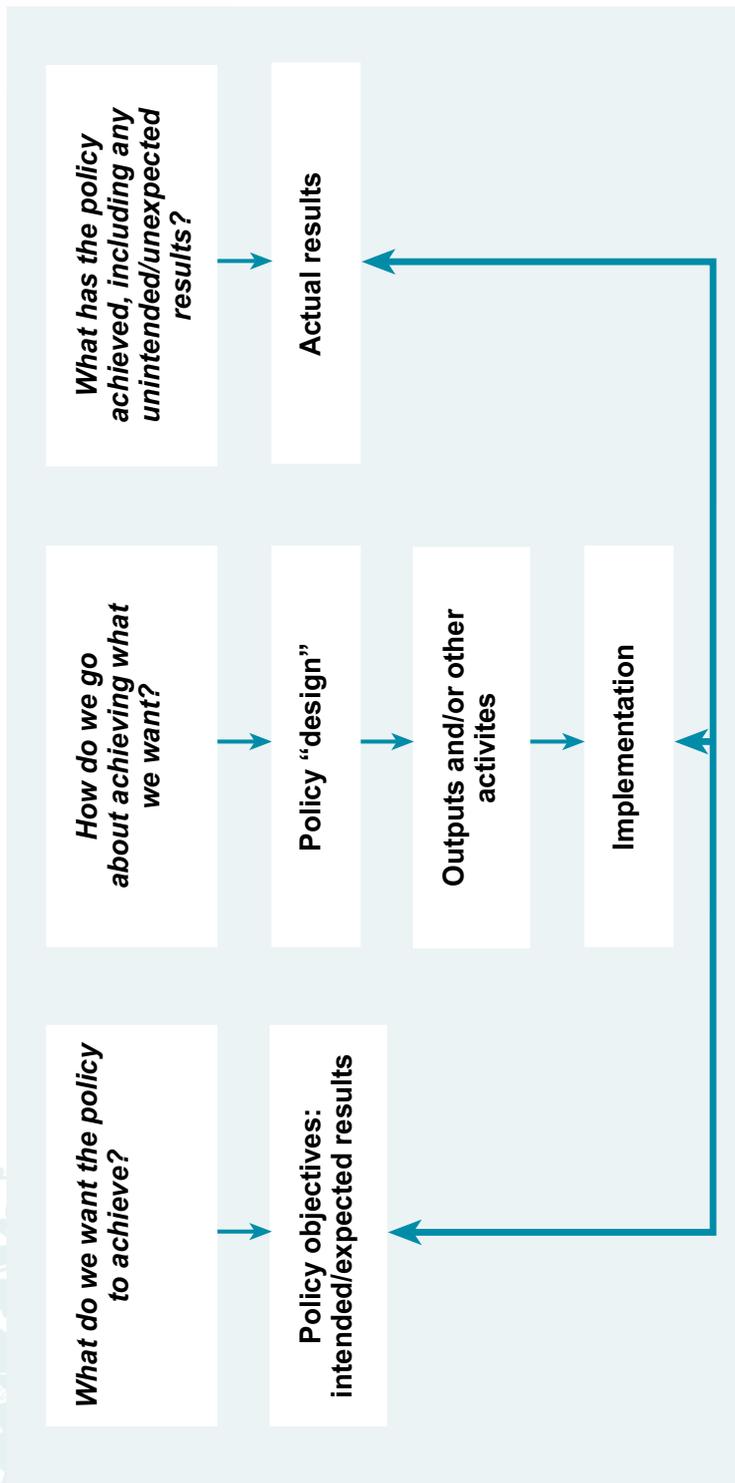
What Is Impact Evaluation?

6.014 Impact evaluation can briefly be described as going about answering the third of the questions in paragraph 6.012 and, at the same time, assessing whether the answer chosen to the second question brought about the actual results.² In practice, however, this is not likely to be a straightforward exercise, because of difficulties in:

1 Parliamentary paper B.29[99c], pages 43-56.

2 An alternative description is determining “what happened” and “how” and the relationship between them.

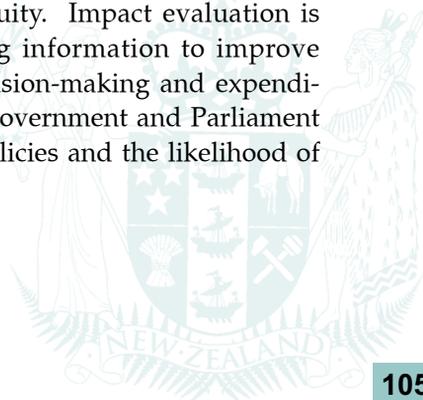
Figure 6.1
A Policy Performance Model



- defining – in objective and quantifiable terms – the policy objectives that you want to achieve;
 - identifying an appropriate and reliable means by which to achieve those objectives; and
 - establishing – with the requisite degree of certainty – causality (i.e. How? and Why?) between the means used and the results achieved.
- 6.015 The third of those difficulties is the defining feature of “impact evaluation”, and is probably the hardest one to resolve.
- 6.016 We canvassed aspects of these difficulties in Part Four of our *Third Report for 1999* and discuss them in more detail in the following sections of this article.

The Place of Impact Evaluation

- 6.017 We have stated previously our belief in the value of impact evaluation of key areas of government spending. In this article we address impact evaluation as a public management tool from the perspective of Ministers and Parliament as decision-makers.
- 6.018 The purpose of impact evaluation is to provide decision-makers with objective, frequently empirically based, information relevant to the decisions they are seeking to take.
- 6.019 Governments commit a significant amount of public money every year on both existing and new policies. All governments want the maximum possible impact from every taxpayer dollar they spend – in terms of efficiency, effectiveness and equity. Impact evaluation is an important tool in providing information to improve the quality of Government decision-making and expenditure, as it helps to inform the Government and Parliament about the success of existing policies and the likelihood of success of policy proposals.



- 6.020 To support its policy decision-making role, and its role in preparing legislative and budget proposals for consideration by Parliament, the Government should have some confidence that the objectives of both ongoing and new policies are achievable.
- 6.021 To support its scrutiny role in passing legislation and making appropriation decisions, Parliament should also have some confidence that both the ongoing and new policies proposed by the Government are likely to be successful.
- 6.022 The information generated through impact evaluation will inform decisions about the design, operation and retention of existing policies, and about the nature and design of new policies. Impact evaluation findings should identify:
- ineffective policy actions that need to be modified in order to achieve the desired outcomes or terminated (during policy implementation or service delivery); and
 - the probability that a new policy proposal will be successful (through the policy formulation process).
- 6.023 Thus, the Government and Parliament should review the continued relevance of existing policies to assess whether their objectives remain relevant to the Government's overall goals, and whether and how well their implementation is contributing to the achievement of the intended policy objectives. Similarly, in determining whether to support new policy proposals, decision makers should ask how they are to know that the proposal is likely to be successful and over what period.



Our Expectations for Effective Impact Evaluation

6.024 We consider that there are a number of pre-conditions for effective impact evaluation at a Government-wide level:

- Policy advice and Government policy decisions will provide a sound basis for subsequently judging whether the policy was effective. Policy decisions will provide an indication of how the Government will know in the future that the policy has been successful and logical criteria for evaluation.
- In order to provide those evaluation criteria, policy advice and decisions will set out –
 - A definition of the problem(s) which the advice is addressing. The problem definition may itself be informed by previous empirical work (where evaluation of similar policies has been completed).
 - The purpose and objectives of the policy in a way that is clearly related to the problem(s) identified. Policy objectives may be statements of desired outcomes at different levels of specificity, outcome targets over different periods, and other expectations.
 - The characteristics of the recommended policy, including the limits of its implementation – such as the characteristics of target groups; financial, age and other thresholds; and regulatory constraints.
 - A soundly based argument for why and how the recommended policy is expected to address the problem(s) successfully – including any critical inter-dependencies and risks, and options for managing both.
- Systematic measurement of critical outcome indicators, and indicators of the implementation of the policy, will occur throughout the life of the policy in order to provide some of the data to be used to assess its success.³

³ Where the policy action is an output, the data may be the same as that reported by agencies and Ministers for accountability purposes.

6.025 Those pre-conditions are based on the assumptions that:

- impact evaluation is a critical source of information for the provision of high-quality policy advice to Ministers and the Government;
- the conduct of impact evaluations is an integral part of, and not additional to, the policy development process;
- not all policies will or necessarily should be evaluated; and
- impact evaluations will usually focus on those policies that –
 - have a significant strategic focus for the Government; or
 - are of a significant cost to the community; or
 - have characteristics that suggest that the continued relevance of the policy is doubtful (even if the amount of public expenditure involved is relatively small); or
 - affect other significant policy areas where there is a need to understand the success of the policy before proceeding with a new policy.

6.026 We expect that each impact evaluation project would:

- Be explicitly assessed as being practicable before the project is undertaken, based on consideration of –
 - whether meaningful data has been or can be measured, meaningful analysis of data is feasible, causal relationships will be discernible through modelling, extrapolation and so on; and
 - the general nature of findings means that they are likely to be useful and able to be acted on by Ministers.
- Be well designed before the project commences, preferably based on the policy decision.
- From the wide range of evaluation methodologies and techniques available, utilise tenable methodologies appropriate to the policy being evaluated. The chosen approach should be free of avoidable biases and should have addressed other ethical considerations.

- Involve analysis of comprehensive, valid and reliable data.
- Be reported to the Government (or an agent, such as a government department) with conclusions that are explicitly derived from the analytical findings.



Our Survey of Impact Evaluation in Use

6.027 This section sets out the results of the survey we conducted to establish the extent to which government departments were using impact evaluation.

What Did We Want to Examine?

6.028 As stated in paragraph 6.015, the defining feature of impact evaluation is the analysis of causal relationships, which itself requires systematic measurement of both aspects of performance being examined – i.e. the action and the impact.

6.029 Therefore, we wanted to establish both whether departments were involved in evaluative activity and the extent to which that activity focused on seeking to establish the impact of Government policies. That is:

- the extent to which departments were measuring, over time, selected dimensions and indicators of both Government policies and their associated outcomes; and
- whether departments were analysing the strength of the link or causal relationship between the two.

How Did We Undertake the Survey?

6.030 We conducted a survey of the 31 central government departments that we considered most likely to be using impact evaluation or other forms of evaluative activity. We provided the departments with our definition of impact evaluation and a summary of our expectations.

6.031 The survey was based on a questionnaire broadly based on the expectations set out in paragraphs 6.024-6.026. We visited each department to discuss the questionnaire before the department completed it and undertook follow-up visits to clarify aspects of the department's responses when asked to do so.

6.032 We asked each department to identify three of the most significant Government policy initiatives taken in relation to their responsibilities over the last five years. For each of these policy initiatives we sought information on whether:

- the policy advice that the department had provided was supported by evaluations of existing or past programmes or empirical research evidence;
- clearly specified, measurable outcomes and policy objectives were articulated as part of the policy decision;
- dimensions of the objectives(s) and the actual outcome(s) had been measured since the decision had begun to be implemented; and
- analysis of the strength of the relationship between the policy design and intended outcome(s) had been undertaken.

What Did the Survey Show?

6.033 In summary, our survey showed that:

- departments are undertaking useful evaluative activity (refer paragraph 6A.005 on page 132), although little impact evaluation;
- departments do not have a clear, common understanding of the nature of impact evaluation (the understanding that departments have is influenced by the nature of their role and functions);
- despite these different starting points, our discussions with departments showed that they support the concept of impact evaluation that we have used;
- desired outcomes and policy objectives are generally poorly specified and therefore provide an inadequate basis for impact evaluation; and
- departments are using a range of evaluation techniques.

6.034 The survey also confirmed concerns about the relatively short times for both Government budget cycles and parliamentary terms compared with the length of time frequently required for the achievement of policy objectives.

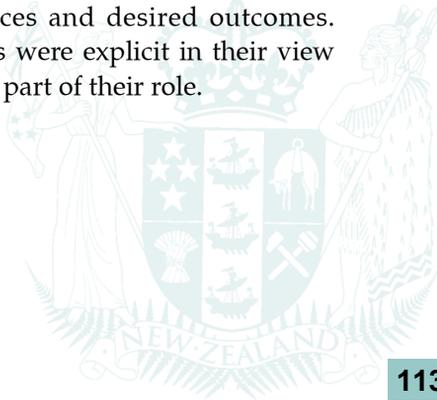
- 6.035 Finally, the survey raised general issues relating to the public management system and (in particular) the requirement of the Public Finance Act for departments to report the link between outputs and outcomes.
- 6.036 Each of these findings is discussed in paragraphs 6.037-6.068.

Departments Are Undertaking Useful Evaluative Activity

- 6.037 We were impressed with the level of general evaluative activity being conducted by departments, and also by the work being done to start systematic outcome measurement. We were able to identify few departments undertaking impact evaluation, but a number of departments are moving towards it – as the two case studies in the final section of this report show.
- 6.038 The main focus of evaluative activity centred on measuring and monitoring aspects of output performance and reviewing delivery methods and processes in order to improve output delivery.
- 6.039 Some departments have been involved in systematic outcome measurement for some time. We were pleased to note that a number of other departments are also beginning to measure aspects of outcome performance.
- 6.040 Generally, departments were conducting evaluative activity in relation to new, small and discrete policy initiatives at the margin of Government expenditure. We did not find any clear indications that impact evaluation was being undertaken in relation to large or strategic policy initiatives.
- 6.041 There appear to be few examples of systematic analysis of the linkages between the outputs delivered and the outcomes achieved, as in impact evaluation. Departments seldom are able to reliably identify the manner in which the implementation of a policy has contributed to outcomes.

There Is No Clear, Common Understanding of Impact Evaluation

- 6.042 Analysis of the questionnaire responses identified that departments held differing views of what constitutes impact evaluation – as opposed to, for instance, research. Some departments regarded impact evaluation as being distinct from, rather than an integral element of, policy analysis.
- 6.043 Some departments considered that measuring dimensions of either output or outcome performance was sufficient on its own, without establishing the linkages between them. Other departments discussed seeking to establish the relationship between policy advice and desired outcomes. This approach creates additional external factors relating to the differences between the advice tendered and the policy decision. The issues that arise in seeking to evaluate the impact of policy advice are discussed further below.
- 6.044 Each department’s view of evaluation and the type and extent of evaluative activity being undertaken was influenced by the role of the department.
- 6.045 As expected, with impact evaluation being integral to policy analysis, policy agencies – in particular sector-based agencies – appeared to have a stronger understanding of impact evaluation. Policy agencies indicated that policy advice was frequently informed by analysis of research findings, and accepted the value of impact evaluation findings as another important source of data.
- 6.046 Service delivery agencies tended to focus their evaluative activities on dimensions of output performance and service delivery processes and did not tend to seek to establish linkages between those services and desired outcomes. Some service delivery agencies were explicit in their view that impact evaluation was not part of their role.



Departments Support the Value of Impact Evaluation

- 6.047 Despite the above differences in starting points, all of the departments surveyed supported our concept of impact evaluation and its potential value for policy analysis and advice and decision-making by the Government and Parliament.
- 6.048 No departments gave technical difficulties or excessive costs in establishing linkages as the reason for not yet undertaking impact evaluation. Some departments commented on cost as a general constraint.
- 6.049 While many policy agencies were at the early stages of designing or conducting impact evaluations, they acknowledged that they still have a considerable way to go to implement impact evaluation as a routine, integral element of policy analysis.

Desired Outcomes and Policy Objectives Are Poorly Specified

- 6.050 Clearly specified, measurable outcomes are not articulated consistently well as part of policy advice and policy decisions. Departments acknowledged that while high-level goals may be specified, more explicit focus was required on the specification of policy objectives and outcome indicators and measures. This information would then provide the basis for examining the success of the policies.
- 6.051 One of the reasons given for inadequate articulation of policy objectives was that policy advice is sometimes tendered directly in response to ministerial direction rather than as the result of a systematic or comprehensive policy formulation process.
- 6.052 A number of departments discussed the inadequacy of the strategic result areas and strategic priorities and overarching goals for analysing the impact of policies. Some departments indicated that, in a general sense, more meaningful policy objectives are sometimes discernible from the objectives specified in legislation. They evaluate compliance with those objectives, although these objectives also tended to be inadequate for assessing impact.

Different Types of Policies Require Different Evaluation Approaches

- 6.053 The inherent nature of some policy initiatives and, to a lesser extent, different policy sectors, means that some outcomes may be easier to measure than others and some causal relationships may be easier to establish. For example, social policy and environment initiatives tend to present more particular difficulties than other policy areas due to the generally long time over which outcomes are expected to be realised. This constraint may be addressed through establishing hierarchies of outcomes with intermediate outcomes to be realised in a shorter time.
- 6.054 A further constraint identified was that policy initiatives are seldom implemented in isolation. This constraint requires careful selection of the evaluation methodology and specific techniques that are appropriate for addressing the particular Government policy.
- 6.055 Departments will sometimes be able to identify clear linkages and a strong causal relationship, but in other situations may rely on less direct indicators and greater use of explicit deduction.

Systemic Issues

Timing

- 6.056 The survey also confirmed a number of time-related weaknesses that are inherent in the public management system.
- 6.057 The one-year Government budget cycle is not conducive to impact evaluation, even with the financial planning period covering three years. Few new policy initiatives can be introduced and fully implemented within one year, and it frequently takes longer for an initiative to reach sufficient maturity to enable any analysis of its impact to be evaluated meaningfully. Evaluations that are carried out during a pilot programme tend to focus more on matters of implementation than impact.

- 6.058 Because of the three-year parliamentary term Ministers frequently seek to establish that they have made a difference and are increasingly requesting evaluations as soon as the policy has been implemented. As discussed above, it is difficult for impact evaluation to be meaningful when full implementation may take a number of years, and the real effects of the policy may not be realised for considerably longer than that.
- 6.059 A further consequence of ministerial expectations is that evaluation findings are not necessarily feeding into decision-making. Frequently, the pressure to move on to the next decision means that advice is tendered and decisions taken before a meaningful impact evaluation of an earlier related decision is possible.
- 6.060 In our view these timing factors lead to ‘evaluations’ being driven towards process reviews and analyses. These evaluative activities are intended to confirm compliance of implementation with design and budget parameters and to improve delivery methods, rather than to identify causal relationships between the policy actions and actual outcomes.

Other Issues Raised by the Survey

Evaluating the Impact of Policy Advice

- 6.061 The survey raised a general issue relating to whether it is meaningful to require a link to be reported between policy advice outputs and the impact of the Government policies that to varying degrees are based on that advice. However, examining such a link is unlikely to be practicable given its tenuous nature and the external factors involved in policy making.
- 6.062 This issue is in even sharper focus when it is considered in relation to policy agencies that have a “second opinion” role, such as population-based policy agencies like Te Puni Kōkiri, or central agencies involved in policy development led by other agencies.

- 6.063 In a technical sense, the desired outcomes of policy advice tendered by a lead policy agency – such as the Ministry of Women’s Affairs – relate to the nature and quality of Government decision-making and the nature of Government policy initiatives. Similarly, the impact of the policy advice provided by service delivery agencies – such as the Department of Work and Income – during the policy development process relates to the extent to which they are successful in persuading the lead agency to tailor the advice to address the particular concerns of the different populations and other interest groups.
- 6.064 In our view, these relationships are not suitable for examination through impact evaluation.

Impact Evaluation and Service Delivery Agencies

- 6.065 The primary incentive of those agencies with a largely service delivery role is to focus their evaluative activities on the delivery of outputs, including output measurement and process reviews. Our survey results indicate that agencies seldom appear to analyse the impact of their services. However, given the closeness of the relationship between impact evaluation and policy analysis, perhaps it is not reasonable to expect service delivery agencies to undertake impact evaluation.
- 6.066 The issue raised by these findings is whether the Public Finance Act requirement to identify the link between outputs and desired outcomes is reasonable or even meaningful in all situations. Greater consideration needs to be given to determining the nature of the information that will satisfy the requirement.
- 6.067 For instance, it may be sufficient for a service delivery agency to assert the link broadly based on the objectives articulated in policy decisions. On the other hand, policy agencies may be required to set out the outcomes that provide the focus of their policy work programme.
- 6.068 Both service delivery agencies and policy agencies could demonstrate in their annual reports to Parliament the links between policies and outcomes based on the findings of impact evaluations completed in (say) the previous three years.

Two Case Studies

- 6.069 In this section we briefly describe two examples of evaluative activity recently carried out by different departments. The purpose of presenting the case studies is to demonstrate that impact evaluation is a practical tool for both policy advisers and policy makers.
- 6.070 The case studies are the evaluations of:
- the Home Detention Pilot Programme of the Department of Corrections (June 1997); and
 - the Supplementary Road Safety Package (LTSA and the Police – July 1998).
- 6.071 The findings from these case studies, in conjunction with other information, have been used to inform further policy advice to Ministers on modifications to the characteristics of the policy initiatives.
- 6.072 Each of the cases clearly indicates progress towards the use of impact evaluation studies. Each study:
- has some characteristics that are consistent with our expectations of what constitutes impact evaluation; and
 - in some respects and to differing degrees, falls short of those expectations.
- 6.073 The completed evaluations tended to focus as much if not more on assessing aspects of delivery than the relationship between the actions and the desired outcomes. Nevertheless, the evaluations provided useful insights and recommendations for modifying aspects of the delivery, at a management level, to better achieve the desired outcomes.
- 6.074 Overall, the case studies point to the progress the government sector is making towards undertaking impact evaluation of the kind discussed in this article.
- 6.075 The case studies do not involve a direct examination of the evaluations themselves, and therefore they do not consider our expectations relating to individual evaluations (paragraph 6.026).

Home Detention Pilot Programme

The Policy Being Evaluated

- 6.076 The Criminal Justice Amendment Act 1993 provided for the establishment of pilot home detention schemes. The Department of Corrections (and before the restructuring the Department of Justice) operated a pilot scheme with the first inmates released to home detention in March 1995. The pilot was to run for two years.
- 6.077 Two types of surveillance were used, designed to provide support and control structures for detainees:
- Passive electronic monitoring of the detainees, involving random telephone calls and a combination of visual and voice verification, in order to confirm compliance with the primary conditions of their release. A home detention officer was available at all times to verify violations recorded by the equipment.
 - A supervisory relationship with each detainee by a home detention officer, including a regime of visits and random telephone calls to the detainee's home and workplace.
- 6.078 The legislation set out criteria for determining the eligibility of inmates to participate in the pilot and the conditions that the detainees must meet while on home detention.
- 6.079 The pilot was designed to cater for a maximum of 30 detainees at any one time, although the actual maximum was 12 with an average of seven.



The Policy Objectives/Desired Outcomes

- 6.080 The reported objectives of the home detention policy were to:
- ease the transition of inmates back into the community through a staged release process and thus to achieve reintegration; and
 - provide home detention as an option only for eligible inmates not otherwise able to be released on parole.
- 6.081 As a reintegrative programme, the home detention policy aimed to reduce reoffending by detainees both during and after the home detention period. However, reducing reoffending was not reported as a formal objective of the home detention policy.
- 6.082 The policy also aimed to release some inmates from prison earlier than would otherwise have been the case, even though eligibility for the programme coincided with eligibility for parole. As home detention was regarded as part of the prison sentence, the pilot sought to avoid releasing inmates on home detention who would have been granted parole.
- 6.083 The evaluation also considered a number of objectives that, while not formally part of the policy objectives, were considered important in assessing the policy's effectiveness. These additional objectives related to:
- the cost-effectiveness of the programme, relative to other forms of imprisonment;
 - compliance with the conditions of the programme by detainees;
 - reducing reoffending by detainees; and
 - minimising the (negative) impact of the home detention programme on families.

Methodology Used for the Evaluation

- 6.084 The evaluation was to cover the first 18 months of the pilot programme and had three parts.
- 6.085 The **overall objectives** addressed the effectiveness of the programme in meeting its objectives (see paragraph 6.080) and whether the programme could be extended to a national system. This part of the evaluation also sought to identify any improvements needed and the features of a national home detention system.
- 6.086 The **process objectives** related to describing the operation of the programme – including the roles and relationships of key personnel, the use of electronic monitoring, and the views of the inmates and their families on the adequacy of the services.
- 6.087 The **outcome objectives** addressed the additional objectives listed in paragraph 6.083. The outcome objectives encompassed:
- describing the rates of successful completion and of reoffending while on home detention;
 - assessing the appropriateness of home detention in terms of the impact on family members and different ethnic groups; and
 - assessing the costs of home detention relative to other forms of imprisonment and parole.
- 6.088 The data gathered during the evaluation was analysed to assess the relationship between the use of home detention and the desired outcome of easing the transition of inmates from prison to the community. Information was gathered from different sources using different methods. The following techniques were used:
- semi-structured interviews (both face-to-face and telephone) with detainees and their families, employers of detainees, and a range of key informants, including Department of Corrections staff (employed in prisons, probation and home detention areas), chairpersons of District Prisons Boards, the contractor for the electronic monitoring element, and prison inmates;

- analysis of case records (including data on rates of successful completions, breaches, reoffending, reviews and recalls to prison), activity logs kept by home detention officers for two separate weeks during the evaluation period, and departmental expenditure records; and
- observation, with home detention officers, prison staff and District Prisons Boards.

6.089 The evaluation was carried out by members of the Criminal Justice Policy Group of the Ministry of Justice and the former Policy and Research Division of the Department of Justice.

Findings of the Evaluation

6.090 In relation to the two policy objectives (paragraph 6.080), the pilot home detention programme was found:

- to be of variable value as a reintegrative programme; and
- not to have a net-widening impact.

6.091 The findings relating to the effectiveness of the pilot as a reintegrative programme primarily addressed issues relating to the impact of home detention on the detainees themselves, their families and their workplaces.

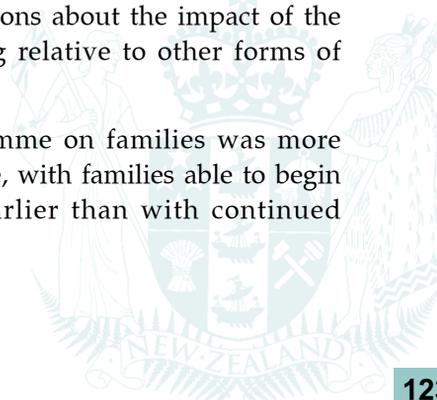
- For the detainees, the findings included improvements in their personal relationships – particularly with their partners, children and parents – improvements in work habits and experience and associated income, and positive comparisons with prison.
- For families, the findings included benefits for family life, relationships and household income, despite some additional burdens.

6.092 The evaluation also reported reservations in the extent to which some of these improvements could be attributed to home detention.

6.093 For all parties involved in the pilot – the detainees, their families and workplaces, and home detention officers – the findings included a number of negative characteristics of

the surveillance aspects of the programme. The legislated requirements for the programme, as it was piloted, concentrated on restricting the detainee – making the predominant focus of the programme one of control rather than rehabilitation.

- 6.094 The findings relating to entry of inmates to the programme were more conclusive, with the results indicating that introducing the programme as a parole option had not led to widespread net-widening. However, the legislated requirements for the programme created little incentive for inmates to agree to home detention since the release conditions were more restrictive than standard parole.
- 6.095 The findings relating to the additional objectives for the evaluation (paragraph 6.083) were that:
- The annual cost for each home detention detainee was calculated to be comparable to the annual cost of minimum-security imprisonment for one inmate. However, this calculation was qualified in that –
 - a number of other factors such as the effects of reduced reoffending could not yet be taken into account; and
 - the programme was operating below its full capacity and, as a pilot, was not able to generate economies of scale.
 - Compliance with the conditions of the programme was satisfactory, within the discretion available to the home detention officers, although some detainees reported rule breaking that was not detected by the monitoring systems. Only one detainee was recalled to prison.
 - The number of detainees on the programme was too low to be able to draw conclusions about the impact of the programme on reoffending relative to other forms of release.
 - The impact of the programme on families was more often positive than negative, with families able to begin restoring relationships earlier than with continued imprisonment.



6.096 The evaluation report went on to make a number of suggestions for improvement, most of which related to:

- improving aspects of delivery through better communication and understanding about all aspects the programme;
- modifications to monitoring equipment; and
- some opportunities for staffing efficiencies.

6.097 The report:

- suggested that it was appropriate to consider the use of an active monitoring system (involving attaching a transmitter to the wrist or ankle of each detainee) rather than the passive system used in the pilot;
- highlighted a number of possible weaknesses in the legislative requirements for both eligibility for, and conditions of, home detention; and
- indicated that a number of the other policy parameters would need to be modified to facilitate reducing prison populations, if that was to become a fundamental policy objective.

6.098 Finally, the report stated that, even with improvements in delivery as discussed above, extending the programme to a national system in the form in which it was piloted would appear to have little purpose.

How the Findings Were Used

6.099 The findings of the evaluation were used, in conjunction with additional information from other local and overseas research evidence, to provide policy advice to the Government on the future of home detention.

6.100 The evaluation also enabled the Department of Corrections to identify areas where improvements in service delivery and performance could be made.

How Did the Evaluation Compare With Our Expectations?

- 6.101 This evaluation met a number of the pre-conditions we expected and clearly indicates progress towards impact evaluation.
- 6.102 Where the objectives of Government policies are not always clearly specified or meaningful the evaluators must establish outcome-related objectives at the start of the evaluation – as in the Home Detention Pilot Programme. (A useful evaluation may still be carried out in this situation. However, it is not ideal as poor policy objectives mean that systematic measurement over time of key aspects of outcome achievement is precluded and information is consequently unavailable to the impact evaluation.)
- 6.103 Most notably, policy advice and policy decisions as presented in the evaluation report included two high-level desired outcomes with no targets. However, the nature of the programme was clearly articulated and systematic measurement of aspects of both the programme and the outcomes occurred – although the report did not explicitly discuss those things.
- 6.104 The evaluation provided useful insights and recommendations for modifying both the characteristics of the programme, at a policy level, and aspects of the delivery of the output, at a management level, to better achieve the desired outcomes.

Supplementary Road Safety Package

The Policy Being Evaluated

- 6.105 The Supplementary Road Safety Package (the Package) was a package of modifications to and extensions of the Government's enforcement and publicity activities, aimed at drink-driving, speeding and seat belt offences.⁴ The Package was to run over four years from 1995-96 to 1998-99.

⁴ The focus on seat belt offences was added to the Package in 1996-97 and was included in the evaluation.

- 6.106 The key actions in the Package were to:
- improve the targeting of speed camera and compulsory breath testing (CBT) enforcement programmes;
 - increase the hours of operation of the speed camera programme; and
 - provide sustained publicity to support the speed camera and CBT programmes.
- 6.107 The Package also involved the introduction of some new equipment:
- advanced laser speed detectors to augment the speed camera programme; and
 - additional breath testing devices to support the CBT programme.
- 6.108 In addition, an independent evaluation of the safety outcomes achieved from the Package was required each year, and modifications to the Package would be made where appropriate.
- 6.109 This evaluation considered the operation of the Package and the outcomes achieved during the two years 1995-96 and 1996-97.
- 6.110 Thus, the Government policy being evaluated was a package representing a combination of modifications to some of the performance characteristics of existing outputs purchased by the Government, new inputs in the form of specific equipment, and management performance through the requirement for regular evaluation and adjustment.

The Policy Objectives / Desired Outcomes

- 6.111 Three critical documents relating to this evaluation provided statements of the Government's policy objectives:
- the National Road Safety Plan 1994-2001 (the Plan);
 - the Safety (Administration) Programme (the Programme) 1995-96 to 1998-99; and
 - the Package.

- 6.112 The Plan set outcome targets – relating to the numbers of persons killed and the numbers of Police-reported injuries on roads in the year 2001– that reflected significant reductions over 1994 levels. The Programme set intermediate targets showing a progressive decrease and also set out targets for a number of behavioural measures (or intermediate outcomes), both of which were considered necessary to achieving the overall outcome targets.
- 6.113 The Package, which was intended to supplement the Programme, set out targets relating to cumulative road trauma reductions in respect of road fatalities, serious injuries and minor injuries.

Methodology Used for the Evaluation

- 6.114 The scope of the evaluation was threefold, with one area of review most pertinent to our exercise on impact evaluation – an assessment of the effectiveness of the Package during 1995-96 and 1996-97. The evaluation also included a review of:
- the evaluation processes established within the Land Transport Safety Authority and other agencies for assessing the effects of the Package; and
 - the implementation of the Package.
- 6.115 The Land Transport Safety Authority contracted independent evaluators (from Australia) to undertake the evaluation.
- 6.116 Completing the assessment part of the evaluation drew heavily on data generated by systematic measurement over time (by different agencies) of the large number of variables required for analysis. These variables included aspects of output delivery performance, environmental characteristics, and outcome achievement.
- 6.117 Quantitative data was available on output delivery, including data on the number of events – such as the number, timing and location of breath screening tests – and on time – such as Police time spent on mobile speed camera activity and driving offences. Data was also routinely collected on the placement and frequency of road safety advertising. Market survey results were available,

providing data that focused on both recall of road safety advertising and views on the likelihood of “being caught” with Police involvement in driving offences.

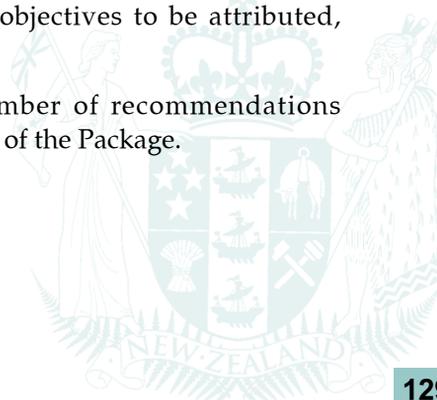
- 6.118 Data was available on the location of accidents, road conditions at the time, driver characteristics and behaviour, vehicle characteristics and condition, and so on. Data was also available on the intermediate and overall outcomes sought through the policy – the number and type of driving offences and the number of road trauma incidents, including details of fatality and injury type.
- 6.119 The primary technique applied during this evaluation was a time-series analysis of serious casualties and fatalities, taking into account characteristics of the Package and factors outside the Package. The analysis took into account the introduction of CBT and speed cameras during 1993, socio-economic factors that may be linked to changes in road use (especially high-risk travel), and trends and seasonal variations in road trauma. For instance, additional time-series analysis was undertaken for serious casualty crashes in terms of both high and low alcohol hours of the week and the location of crashes – urban and rural.
- 6.120 While some factors may have an effect on the outcomes sought, it is not always feasible to include them in the analysis. In particular, it was also acknowledged that although on-going improvements to roads have made a contribution to the downward trend in road trauma, their gradually increasing effect (relatively small change from year to year) made them unsuitable for explicit consideration in the analysis.

Findings of the Evaluation

- 6.121 The evaluation report stated that it had been possible to estimate the effectiveness of the Package and to comment on the contributions of its separate elements. A number of findings were reached about the relationship between the aspects of the Package and the level and type of road trauma and estimates were made of the savings in road fatalities and injuries that could be associated with the Package. Overall, the evaluation concluded that it was likely that the

targets for the reduction in fatalities and injuries during the four years of the Programme had already been met or exceeded in its first two years.

- 6.122 The degree of causal relationship was qualified through reporting a number of indicators that suggested that part of the reduction in the risk of death or serious injury on the roads during 1995-96 and 1996-97 could be attributed to the components of the Package. The findings were graduated with:
- the drink-driving component of the Package being described as suggesting a substantial contribution;
 - the speeding component as suggesting a smaller contribution; and
 - the seat belt component as probably suggesting some contribution in 1996-97 only.
- 6.123 The evaluation also concluded that *the procedures established by the Land Transport Safety Authority for monitoring and evaluating the outcomes of the Supplementary Road Safety Package . . . allow the effects of the SRSP to be seen in terms of changes of advertising awareness, public attitudes, on-road behaviours, and road trauma levels related to drink-driving and speeding.*
- 6.124 The report recommended a number of additional policy initiatives and analysis of existing programmes, with specific reference to speed cameras.
- 6.125 One recommendation related to the development of a mathematical model. The discussion indicated that use of such as model would enable any causal relationship between the components of the Package and achievement of the Government's policy objectives to be attributed, rather than estimated as now.
- 6.126 The report also made a number of recommendations relating to the implementation of the Package.



How the Findings Were Used

- 6.127 The findings from the evaluation of the Package were used by:
- the Ministry of Transport in providing policy advice to the Government concerning the continuation of the Package;
 - the LTSA and the Police; and
 - by other agencies in modifying aspects of the delivery of the outputs for which they are responsible.

How Did the Evaluation Compare With Our Expectations?

- 6.128 This was an impact evaluation that provided a practical tool for both policy advisers and policy makers. The pre-conditions we expected were sufficiently in place through the Government policy decision.
- 6.129 For the Package, the policy advice and policy decisions included clear, specific articulation of both the desired outcomes and the nature of the Package; and the agencies undertook systematic measurement of aspects of both the actions and the outcomes. Outcomes were articulated at several levels of specificity, with:
- ultimate targets for achievement by the end of seven years;
 - intermediate targets for the intervening years; and
 - an additional set of intermediate outcomes and targets that were argued as being necessary to the achievement of the overall outcome.

- 6.130 The evaluation also provided useful insights and recommendations for modifying both the characteristics of the actions within the Package (at a policy level) and aspects of the delivery of those outputs (at a management level) to better achieve the desired outcomes.

Appendix A

Impact Evaluation in the Public Management System

6A.001 The purpose of this section is to describe the place of impact evaluation in the public management system. In conjunction with our expectations relating to its use within the system (paragraphs 6.024-6.026), the section provides the basis for our survey (paragraphs 6.027-6.068) and a framework for considering the two case studies reported (paragraphs 6.069 to 6.130).

The Significance of Impact Evaluation

- 6A.002 Officials are increasingly acknowledging the importance of evaluation as a public management tool. There are numerous conferences on evaluation each year and papers by various government agencies and commentators.⁵
- 6A.003 However, agreement has not necessarily been reached about the nature (definition) or characteristics of good evaluation, or even the language of evaluation. Nor has there been systematic or consistent use of evaluation practices.
- 6A.004 The topic is vast and this article is neither a general exploration and discussion on the various approaches to evaluation used by government agencies nor a theoretical or conceptual paper on definitions of and differences between various types of evaluation or evaluation methodologies.

⁵ For instance, *Looping the Loop: Evaluating Outcomes and Other Risky Feats*, State Services Commission (1999).

- 6A.005 We use the term “evaluative activity” for describing numerous situations in which government agencies measure, monitor, review and analyse aspects of entity performance – including outcome achievement or status, output delivery, and input use. These activities do not constitute impact evaluation.⁶
- 6A.006 Impact evaluation findings are generated through two stages:
- First, systematic measurement over time of selected dimensions and indicators of both policy actions – primarily outputs – and their associated outcomes.
 - Secondly, analysis of the causal relationships between the two, which is the critical characteristic of impact evaluation.
- 6A.007 Impact evaluation is not limited to measuring and reporting outcome achievement, output delivery or the operation of other (non-output) Government policy actions.
- 6A.008 However, evaluation findings will seldom demonstrate causality conclusively, and the strength of causal relationships that can be demonstrated will be weaker in some situations than others. Systematic measurement of the characteristics to be studied and comprehensive analysis of trends and patterns is required to be able to attribute changes to a particular factor with sufficient confidence for the findings to be meaningful.
- 6A.009 These limitations are recognised and accepted characteristics of impact evaluation. Nevertheless, findings of reliable impact evaluations will always provide a more objective and higher-quality platform of information on which:
- policy advice may be tendered by departments;
 - policy decisions may be taken by Ministers and the Government; and
 - legislative decisions may be taken by Parliament.

⁶ Evaluative activities include operational audits, performance auditing, reviews, customer satisfaction surveys, and routine measurement of aspects of output delivery.

Impact Evaluation and Policy Advice

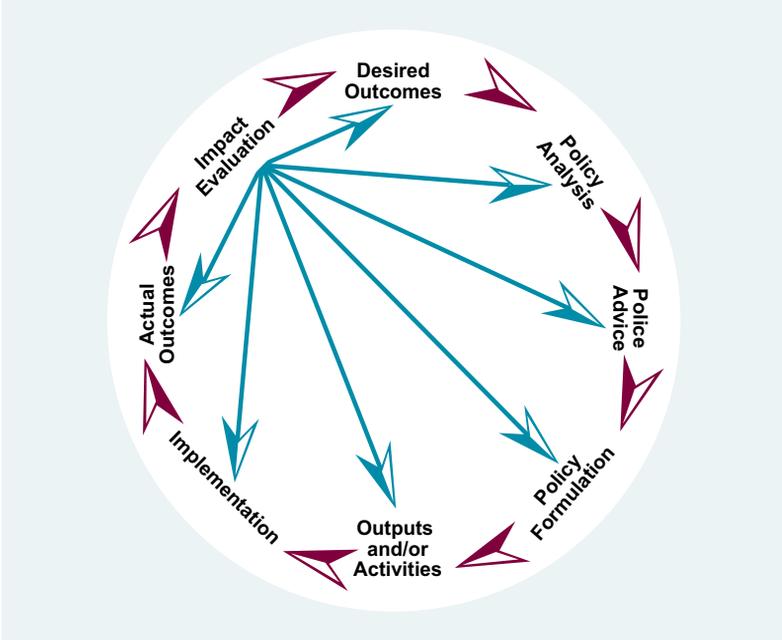
Policy Agencies

- 6A.010 Policy agencies are responsible for providing policy advice to their Ministers and the Government and, as an inherent element of that advice, for presenting a view on the nature of the causal relationship between different policy design options and the Government’s desired outcome(s).⁷
- 6A.011 Greater use of empirically based policy advice is likely to inform higher-quality decision making – and a movement away from largely deductive approaches to policy advice. Thus, in order to be high-quality policy advice, that advice should be informed by (among other things) current information and empirical data on the effectiveness of related existing Government policies and how those policies have affected achievement of the outcomes the Government is seeking.
- 6A.012 Impact evaluation is an important source of such data. Impact evaluation is also an important source of information to determine the level of consistency between the design of a policy and the way in which it has been implemented.
- 6A.013 The place of impact evaluation in the policy circle is illustrated in Figure 6.2 on page 134.

⁷ Limitations on impact evaluation arguably also apply to policy advice – see paragraphs 6.061-6.064.



Figure 6.2
The Policy Circle



6A.014 Evaluation during formulation of a new policy is likely to be a desk-based exercise using empirically based findings from evaluations of similar policies as well as relevant research material. Using these critical sources of information advisers extrapolate conclusions about the likely success of different options for addressing a particular policy problem.

6A.015 In this way, evaluation findings are used to give shape to policy actions and assist decision-makers to determine the “right things to do” to address particular policy problems. Evaluation findings also assist decision-makers to determine the general design of the policy actions and the parameters under which they will be implemented.

6A.016 Impact evaluation during policy implementation will also consider whether the findings demonstrate that the policy design is the “right thing” for achieving the desired objectives. In addition, impact evaluation at this stage will consider whether any differences between the policy design itself and implementation of the design have improved or lessened the effectiveness of the policy.

Service Delivery Agencies

6A.017 On the other hand, service delivery agencies are responsible for optimising the achievement of desired outcomes within the policy design and other parameters agreed by the Government, through the selection of service delivery methods. Fulfilling this responsibility is helped by ongoing monitoring of the implementation in order to inform operational decisions about those aspects of a policy that are (reasonably) controllable by the service delivery agency.

6A.018 While being important for management purposes, evaluative activities of that type do not constitute impact evaluation as discussed here. Such monitoring considers whether implementation of the policy is consistent with the policy design and may identify ways in which altering the service delivery methods or approaches (consistent with the policy design) may improve the effectiveness of the policy. Such evaluative activities will identify whether the agency is “doing things the right way”.

The Focus and Findings of Impact Evaluation

6A.019 The focus of an impact evaluation will depend on the nature of the policy question being considered and the purpose of the information that is sought. Similarly, the methodology adopted for an impact evaluation will depend on the nature of the policy being examined.

6A.020 Impact evaluations may, for instance, seek to identify:

- the actual effect of a particular Government policy; or
- the policies that are contributing to a particular outcome area.

- 6A.021 Thus, impact evaluation may relate to a complex umbrella Government policy involving a number of policy initiatives expressed through several appropriations and a number of Votes or departments. Impact evaluations may also relate to a specific policy expressed through a single or multiple appropriation and single or multiple Votes or departments.
- 6A.022 Evaluation findings may relate to:
- the design of the policy;
 - the manner in which the policy has been implemented, especially where there may be some differences between implementation and policy design; or
 - the definition of the problem which the policy was intended to address.
- 6A.023 Findings may be based on explicitly stated logic (deductive reasoning) – supported by a correlation between data sets, data modelling and extrapolation – and other analytical techniques.
- 6A.024 Thus, impact evaluation is a critical and systematic analysis, using empirical data, of whether the results intended by a policy have been or are being achieved (for ongoing expenditure) or are likely to be achieved (for policy and expenditure proposals). Impact evaluation may also indicate whether the design of current policies may need to be changed to better contribute to desired outcomes.



Appendix B

The Current Legislative Framework

6B.001 The Public Finance Act 1989 (the Act) reflects the first two elements in our policy performance model (Figure 6.1 on page 104) by:

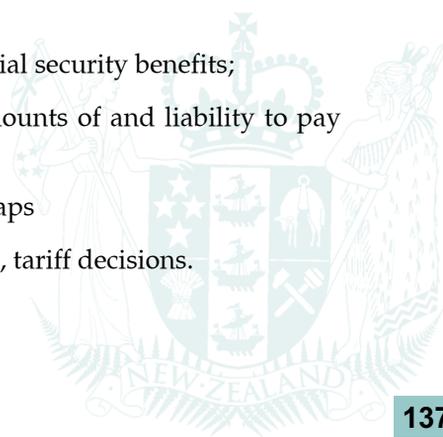
- defining “outcomes” as *the impacts on, or consequences for, the community of the outputs or activities of the Government*; and
- requiring the *Estimates of Appropriations* to *identify the link between the classes of outputs to be purchased by the Crown and the Government’s desired outcomes*.

6B.002 However, what the Act does not do is to provide any guidance on:

- how to describe outcomes, with associated measures or criteria, in such a way as to be able to establish whether or not they have in fact occurred; and
- how strongly the “link” between outputs and outcomes should be identified.

6B.003 Furthermore, the Act does not explain what is meant by “activities” (as something different to outputs) of the Government, nor does it require that any link between “activities” and outcomes be identified. We assume that the term embraces such significant aspects of Government fiscal measures as:

- transfer payments, e.g., social security benefits;
- revenue decisions, e.g., amounts of and liability to pay taxes;
- capital spending; and perhaps
- non-budgetary actions, e.g., tariff decisions.



- 6B.004 Nevertheless, probably the greatest omission from the Act is any explicit requirement to report what outcomes have occurred, with an explanation of how they compared with the intended outcomes. (It is possible to interpret the Act as inferring such a requirement through the obligation to produce statements of objectives and statements of service performance. However, those statements are directed only at outputs.)
- 6B.005 Thus, the extent to which our policy performance model is reflected in the Act can be represented as shown in Figure 6.3 on page 140.

Difficulty In Defining Outcomes

- 6B.006 There are a number of inherent hurdles for governments in articulating desired outcomes, meaningful policy objectives, and (especially) targets for the achievement of outcomes. The primary hurdles are the knowledge that in reality many factors contribute to the achievement of outcomes – only some of which are within the reasonable control of any government – and that outcomes are generally achieved over periods longer than parliamentary terms.
- 6B.007 Outcome specification – and subsequently impact evaluation as a practical tool – should properly be limited to outcome targets that indicate progress towards the desired outcomes and policy objectives that are intermediate to the achievement of high-level outcomes.⁸ These approaches are likely to address any concerns about the time required for the achievement of outcomes.
- 6B.008 For instance, a health policy decision may include purchase of both education and regulatory services. The objectives of this policy, the desired outcomes, may relate to reducing the level of tobacco-related disease and deaths. The relationship between these services and the incidence of disease over time could be the subject of an impact evaluation.

⁸ Other forms of evaluative activity, such as social science research, which are better placed to address these higher level questions, are outside the scope of this study.

- 6B.009 Progress towards the achievement of the outcomes may be measurable over a relatively short time. A correlation is likely to be able to be demonstrated between changing levels of disease-causing behaviour in that period and the policy actions, once external factors such as other known causes of the same diseases have been taken into account.
- 6B.010 It is unlikely that an impact evaluation would be able to indicate a clear causal relationship between a policy such as this example and the health status of the general population as compared with the smokers in the population.



Figure 6.3
The Public Finance Act Policy Performance Model

