

# Policy advice as a process; competency and evaluation issues

by

Robin Johnson<sup>1</sup>

*There is an increasing interest in and demand for policy advice which articulates desired outcomes and explains how the advisory and operational outputs of a ministry will contribute to advancing these. Increasingly, ministries are being challenged, for example by the Audit Office and select committees, to provide coherent definition and explanation of the "strategic outcomes" and the means to achieve them (Chapman 1995).*

*We are convinced that more research and evaluation of key areas of government spending would be extremely beneficial. Properly directed and conducted, such studies would yield valuable information to support policy development. They would also reduce the risk of important outcomes not being achieved and substantial sums of public money not being wasted (Office of Auditor-General 1999).*

## Introduction

There is a current perception that more 'research and evaluation' would be beneficial to the quality of policy advice in the government system. While evaluation is used to mean measuring the longer term effects/results/impacts of particular policies, which will necessarily involve detailed research, monitoring is included here to represent ongoing measurement to give feedback to sponsors and/or policy decision makers<sup>2</sup>. More of one service usually means less of another so this issue is of considerable importance to decision makers in government. If more resources are devoted to monitoring and evaluation, will there be a payoff? What is the balance between asking departments to devote more resources to internal and external research into policy issues and pursuing other activities deemed relevant by Ministers?

Discussion of this topic revolves around the pursuit of outcomes and outputs in the government system. Outcomes are those broad categories of government policy objectives that define what the political arm of government is trying to achieve. Outputs are categories of government services, on the other hand, which are produced by departments ostensibly to meet the outcomes desired by Ministers. These concepts

---

<sup>1</sup> Formerly Assistant Director and Policy Consultant, Economics Division, Ministry of Agriculture and Forestry (johnsonr@clear.net.nz). I have benefitted from discussions with Robyn Reynolds and Mark Robinson (SSC), David Galt (Treasury) and Bryce Wilkinson; but they are not responsible for what I have written here.

<sup>2</sup> This is World Bank usage of the terminology. See World Bank (1998), *Assessing Development Effectiveness*, Operations Evaluation Department, Washington, p.28.

were developed in the discussions on departmental performance following the Public Finance Act 1989<sup>3</sup>

The debate about outcomes and outputs has obscured the underlying problem with regard to outputs, i.e. with regard to the efficiency and effectiveness of policy measures carried out in the past and proposed for the future. Monitoring is concerned with measuring progress with a policy initiative or project in regard to some laid-down set of objectives or achievements, while evaluation is concerned with making judgements about both efficiency in the use of resources and effectiveness in terms of some laid-down goals or the wider goals of government.

The government has not pursued a goal of compulsory evaluation of all proposals submitted for Budget allocations, though the requirement for Regulatory Impact Statements suggests such a goal.. It appears that departments are left to decide on their own approaches for support for budgetary matters subject to Treasury guidelines.. This could lead to an uneven application of evaluation methodology and detract from its overall worth. Nevertheless, it is also true that there is no agreement on appropriate methodologies, hence the possible requirement of compulsion may well be out of reach.

This paper [and the following papers] is concerned with the process of policy advice in the Government system and how effective it is. It is concerned with the role of monitoring and evaluation within the policy advice system. The following discussion indicates that there is political sensitivity on the one hand about revealing the effects of past and current policies, while on the other, there is a contrary movement emphasising that the quality of policy advice could be improved by better background research and communication of results to key policy advisors and thence to key decision makers.

The main burden of monitoring and evaluation of day-to-day impacts of policy change rests with the bureaucracy though outside agencies might be employed in some cases. Central to the discussion is the distinction between outcomes and outputs.

### **Outcomes v Outputs**

The Public Finance Act 1989 established the current framework for purchase agreements between Ministers and CEOs which in turn identified outcomes as the set of broad objectives which Ministers wished to support, and outputs as the goods and services departments deliver through the purchase agreement. There is an understanding that in a broad sense Ministers are ultimately judged on outcomes, while departments/CEOs are held accountable for the production of outputs<sup>4</sup>. The distinction in interests was intended to allow clear demarcation of the lines of responsibility which could be written into CEO contracts. Better performance of the bureaucracy was the end objective.

"Policy advice" is the common output to all departments, among many other particular services provided. In the purchase agreements, this output describes the area of policy the department is responsible for and the types and quantity of advice they are going to provide. Outcomes, on the other hand, are thought to be too broad for contract

---

<sup>3</sup> See G. Scott (2001), *Public Management in New Zealand*, NZ Business Roundtable, pp.176-7.

<sup>4</sup> Scott (2001), op cit, p.175.

specification. They are more like "goals" in the old terminology, which were seen as motherhood type statements which were vague but put across a good feeling suitable for political purposes. In a background document<sup>5</sup> the SSC saw difficulties in measurement and causality as being the reason why departments could not be held responsible for outcomes, and therefore why outputs and not outcomes should be the focus for departments.

The SSC suggested that there had been too great a focus on outputs [that is, what departments were doing and how they were doing], and that departments were starting to lose sight of the outcomes Ministers desired. Outcomes should be more formally incorporated as strategic drivers in the public management system. "A renewed focus on outcomes and in particular more evaluation of the impacts of policy outputs on outcomes should not detract from the current output-based performance management system....*Both ex ante identification of the impacts of outputs on outcomes and the ex post evaluation of the same, could be seen as a complement to the current accountability system.* Instead of getting tied up in debates about accountability we should be focussing on the real goal, finding mechanisms to *improve the quality of policy advice to Ministers*"<sup>6</sup> (italics added).

### **Treasury Procedures and Regulatory Policy**

Treasury is the critical control department on budget matters. The department advises the Minister of Finance on whether the government's spending programmes represent value for money in relation to both the government's objectives and the national interest. Treasury is the control gate through which all policy matters with financial implications must pass.

Documentation is provided to departments by Treasury so that they may meet what the Treasury sees as its requirements<sup>7</sup>. Among the guidelines are clear directives to define what outcome is being sought, whether all the options have been considered (including the status quo), and what are the assumptions behind the declared costs and benefits (where they have been identified)? They ask whether there are clear criteria for analysing the options, and whether the criteria include effectiveness and efficiency considerations?

The key instruction concerns the implementation of the preferred option. Is it clear how ...the option will be monitored and evaluated?<sup>8</sup> What are the likely obstacles to success? How might they be circumvented and what fall-back positions have been considered? Does the paper consider different implementation options(e.g. piloting) where there is insufficient information to roll out the favoured option? Finally, does the paper recommend an evaluation of agreed scope [to] be reported back at some future date?

---

<sup>5</sup> State Services Commission (1999b), Essential Ingredients : Improving the Quality of Policy Advice.

<sup>6</sup> SSC (1999a), Looping the Loop: Evaluating Outcomes and Other Risky Feats, p.10.

<sup>7</sup> Treasury (2001), How Treasury Approaches Draft Papers for Ministers, and Treasury (1999) Guidelines for Costing Policy Proposals.

<sup>8</sup> There is a need to get evaluation of both efficiency and effectiveness built into ongoing Government activities, while avoiding it becoming a compliance activity, with no bite, ensuring appropriate but not excessive resourcing and ensuring the evaluation effort is directed to the highest priorities (D.Galt, pers com).

These requirements ask that outcomes be clearly defined and by implication that departmental outputs are produced in a least cost framework. But Treasury appears to be mindful that excessive resourcing is not desirable and departments choose where evaluation is most cost-effective.

### **Regulatory policy**

For some years, all policy proposals submitted to Cabinet which result in government bills or statutory regulations must be accompanied by a Regulatory Impact Statement (RIS). The Statement should consistently examine potential impacts arising from government action and communicate the information to decision-makers. Completion should provide an assurance that new or amended regulatory proposals are subject to proper analysis and scrutiny as to their necessity, efficiency, and net impact on community welfare. An RIS should cover the need for government action, the public policy objective, a statement of the feasible options, a statement of the net benefits, and a statement of the consultative program undertaken<sup>9</sup>.

This procedure is copied from an Australian model where performance has been monitored by the Productivity Commission<sup>10</sup>. Since April 2001, a Business Compliance Costs Statement has been added to this requirement. Recently, the Minister of Commerce stated that "To be quite frank they [RIPs] were not being used as effectively as they should be. They were a box to tick. They were bland. And nobody took any notice of them"<sup>11</sup>.

This observation raises issues of incentives for completing more appropriate Statements, training for staff involved, and sanctions for non-compliance in producing mandatory departmental outputs..

### **Social Policy Evaluation Developments**

In the social policy area, recent research has been focussed on refining the meaning given to social outcomes<sup>12</sup>. The Strategic Policy Group of the then Ministry of Social Policy has been working on the conceptual foundations of cross sectoral social policy and social development and putting them into practice. "Historically social policy was organised around functional interventions such as economic, education, housing and health policies. The balkanised nature of this policy advice risks poorly coordinated and integrated policy. Policy advisors may fail to analyse either the positive or negative effects of proposed policies in other sectors. The proposed framework for cross sectoral social policy is aimed to overcome this problem and improve the overall coherence of policy analysis"<sup>13</sup>.

"The framework for cross sectoral social policy provides a conceptual model to structure thinking about policy. Such an approach requires:

- a clear articulation of the desirable social *outcomes* that are the goals of policy;
- a strategy for *measuring* these outcomes; and

<sup>9</sup> Ministry of Commerce (1998), A Guide to preparing Regulatory Impact Statements.

<sup>10</sup> Productivity Commission (1998), Regulation and its Review, Ausinfo, Canberra.

<sup>11</sup> *The Independent* Feb 6, 2002

<sup>12</sup> Proctor R. (2001), An Inclusive Economy; Rea D. (2001), The Social Development Approach.

<sup>13</sup> Rea (2001), op cit p.1.

- consideration of how government action *impacts* on all these outcomes both directly and indirectly, and especially over the long term" (italics added).

On the organisation side, MSD convenes the Social Policy Evaluation and Research (SPEaR) Committee, an officials coordinating group set up in late 2001. This recognises that high quality social research and evaluation "has a fundamental role" in the development of evidence-based social policy<sup>14</sup>. Greater coordination of the research spend by social policy agencies is expected to lead to improvements in the uptake of research and evaluation information into social policy development. The role of SPEaR is to oversee the Government's social policy research purchase. In particular, the identification and dissemination of key social policy knowledge needs;

- development of research and evaluation agendas to meet identified social policy knowledge gaps;
- assurance of quality and relevance of Government purchased social policy research and evaluation;
- integration of research and evaluation information into evidence-based policy development; and
- evaluation of the success of recent initiatives to enhance the knowledge base for social policy.

Significantly, MSD is organising a public conference with the objectives of improving effective communication between policy officials and (both departmental and external) research providers about what information is required for social policy development. As a result of wider consultation, they believe there is a lack of complementarity between research programmes and social policy agencies' knowledge needs. Among the primary objectives of the conference is "facilitating presentations of social research and evaluation that have potential for use by social policy agencies and service providers"<sup>15</sup>..

This is clearly a greater thrust toward increased monitoring and evaluation for outcomes in the social policy area. There appears to be considerable emphasis that impact evaluation will be provided by outside agencies. MoRST and the Royal Society are also involved in this initiative as well as university-based people. It remains to be seen how it will impinge on policy advisors and policy makers.

### **The need for incentives and/or sanctions**

In 1999 the SSC view was that there has been a neglect of the evaluation side of the policy process in the past<sup>16</sup>. There was not a formal requirement in the cabinet manuals to carry out such evaluation. Treasury requirements were apparently not explicit enough. The problem was recognised but not made mandatory.

The SSC took the view that a good deal of evaluation already takes place. But most is focussed on evaluation for the purpose of better delivering and implementing of programmes. Less emphasis had been placed on evaluating the impact of interventions

---

<sup>14</sup> MSD website 'SPEaR'

<sup>15</sup> MSD website 'conference objectives'.

<sup>16</sup> For a penetrating discussion of evaluation in the Aid Division of the Ministry of Foreign Affairs and Trade, see *Toward Excellence in Aid Delivery*, Report of the Ministerial Review Team.

on broader outcomes or on how departmental activities contributed to Government's stated policy priorities<sup>17</sup>. Evaluation was typically not built into the policy document at the outset, thereby making future review problematic.

Reasons why evaluation of outcomes was not a strong feature in the NZ context were summarised as:

- low perceived demand ("ministers not interested", information has an upside and a downside);
- fluctuating policy settings (rapid changes in policy settings at elections and other times);
- funding (reflecting lack of priority in the past);
- methodological issues (lack of clear cut answers in social research);
- dearth of skills and capabilities (lack of competent evaluators); and
- self interested stakeholders (ministers and public servants' unwillingness to be scrutinised, each for their own reasons)<sup>18</sup>.

The Public Finance Act has clearly raised the profile of monitoring and evaluation in government processes though decision makers have not moved to the ultimate sanction of compulsory reporting<sup>19</sup>. The inherent logic of the outcome/output distinction means that departments must be able to justify the outputs they put forward for ministerial agreement and budgetary approval. One Treasury official defines the relationship in these terms "Assessments need to be drawn about the relationship of inputs to outputs (technical efficiency or value for money), outputs to outcomes (allocative efficiency or effectiveness) and on changes in departmental capability (physical or intangible investments).....An intention of the reforms was to create greater incentives for departments to assess and reveal performance in each of these dimensions. The spur would come from Ministers acting as a discerning customer. Ineffective outputs would be cut out, and if prices were too high other suppliers would be sought, or if this was not possible changes might be sought to management"<sup>20</sup>.

"To date this goal of departments proving their performance has only been partly achieved. One reason has been the time needed to make and embed changes of this magnitude: lifting the levels of technical and management skills, introducing new systems of funding, reporting, and performance measurement. Changing departmental cultures takes years. Another reason is that greater onus could have been placed on departmental managers to prove their efficiency and effectiveness. Placing the onus of proof onto spending proponents would have increased the incentives on departmental Ministers to seek information from their departments"<sup>21</sup>.

The SSC view of these matters was that existing incentives in the system tended to cause managers to avoid open scrutiny of the relative merits of existing programmes; that the budget focus on new initiatives avoided zero-based evaluation of past

---

<sup>17</sup> SSC (1999a), op cit.

<sup>18</sup> SSC (1999a), pp.6-7.

<sup>19</sup> The Australians tried such a system in the early 1990s. See Di Francesco (1998), The Measure of Policy: Evaluating the Evaluation Strategy as an Instrument for Budgetary Control, *Australian Journal of Public Administration* 57, 33-48.

<sup>20</sup> Bushnell P. (1998), p.2, Does Evaluation of Policies Matter?, Foreign and Commonwealth Economic Advisors, London.

<sup>21</sup> op cit, p.2

programmes; and that departments tended to protect their vote even when another department's programme was found to be more effective. In addition the short-terminism in the system is not conducive to outcome evaluation. Evaluations – especially in the social policy area - typically require a longer time-frame or some ongoing commitment to monitoring progress over time<sup>22</sup>. Sometimes policy formation cannot wait for the necessary research to be carried out. These are powerful disincentives and they will require considerable encouragement and direction from CEOs and Ministers if they are to be reversed.

### *The Audit Approach*

In a review of progress since the passing of the Public Finance Act 1989, the Controller and Auditor-General<sup>23</sup> has discussed what the 'outcomes' mean and how they are described and how these relate to strategic priorities and overarching goals. Successive governments may have been reluctant to be more specific about desired outcomes to avoid the complications that might arise if they failed to achieve them. It was inescapable that if outcome statements continued to be expressed in a way that was vague and unmeasurable, the information conveyed to Parliament was extremely limited and virtually worthless. Could impact evaluations be the answer?.

"Impact evaluations are empirical studies that are conducted to measure and establish the real consequences of an agency's actions or programmes. Unlike some other countries, Parliament had not previously required the impact of government spending be subject to empirical research or evaluation<sup>24</sup>. The absence of empirical information meant that it was difficult for Parliament to obtain assurance from the very considerable sums spent on many government activities that they are having the intended effect (or indeed, any effect)...It is worth noting, however, that impact evaluations are more useful for informing policy decisions than as a tool of accountability. Usually they cannot be undertaken until the programme being evaluated has been in place for sufficient time for the asserted benefits to be realised".

Difficulties in reporting may have lead to lower standards. The key issue is how public entities set direction, and measure and report on performance. "Statutory compliance requires reporting on an entity-by-entity basis. However, a reporting entity as defined by statute will not necessarily be the best level for reporting performance, and sometimes defining a useful reporting entity is difficult. Performance can also be forecast, managed, or reported at different levels of the public sector such as:

- outputs provided - that is, goods and services delivered;
- a set of outcomes which together make up a programme of service delivery;
- the efforts of a single public sector entity;
- the combined efforts of entities within a sector;
- efforts of entities across several sectors to contribute to one policy initiative; or
- the whole of government"<sup>25</sup>.

---

<sup>22</sup> SSC (1999a), p.11.

<sup>23</sup> Office of the Controller and Auditor-General (1999), Third Report for 1999: The Accountability of Executive Government to Parliament (website).

<sup>24</sup> OAG, op cit, p.52.

<sup>25</sup> OAG 2001, Reporting Public Sector Performance, p.15 (website).

## Content and Guidelines

The Public Finance Act established the outcome/output framework for accountability in the policy delivery system but did not set a procedure for evaluation of past policies as part of the policy advice process. A 1995 assessment<sup>26</sup> drew attention to the weak connection between outputs and outcomes. Poor quality policy analysis and solutions were being offered to Ministers as a result. Key areas of difficulty were:

- lack of agreement over outcome specification;
- poor problem definition;
- lack of robustness of causal theory;
- poor empirical information; and
- inadequate review and evaluation of policies.

Treasury confirmed this in 1998<sup>27</sup>, "To date this goal of departments proving their performance has only been partly achieved. One reason has been the time needed to make and embed changes of this magnitude: lifting the levels of technical and management skills, introducing new systems of funding, reporting, and performance measurement".

These prescriptions therefore suggest that better policy could result from greater study of policy programmes in an *ex post* framework. The question of *ex ante* analytical constructs is still a policy requirement for policy advisors within departments, but if lessons are to be learnt from previous experience, then an *ex post* framework is required that is reasonably universal and which can be adapted to the job in hand<sup>28</sup>.

There are variations between the agencies but the broad thrust of impact evaluation strategies is to review efficiency and effectiveness. Efficiency is seen to be related to policy 'outputs' and how they are produced with the resources available, while effectiveness is related to policy 'outcomes' and whether they are achieved. What is not clear is how such evaluations are going to be carried out, what kind of staff is needed, and what training programmes should be provided.

## Multidisciplinarity and training

The disciplines involved in evaluation are diverse. This has wide implications for staffing and training in departments. What are the appropriate disciplines?<sup>29</sup> Discussion of evaluation as an instrument of budgetary control clearly points to economic management. But policy *outcomes* are clearly of a more general nature than economic management and will consequently require a wide range of skills to meet different departmental responsibilities.

---

<sup>26</sup> Chapman R. (1995), Improving Policy Advice Processes: Strategic Outcome Definition, Policy Evaluation and a Competitive Policy Market, *Public Sector* 18, 16-21.

<sup>27</sup> Bushnell (1998) op cit

<sup>28</sup> Scott (2001), p.334, stresses that strategic policy analysis rests on professional analysis of the highest calibre; no ministry or department [in NZ] reaches this standard all the time.

<sup>29</sup> In a discussion of the membership of the American Evaluation Association, the observation is made that its members were mostly trained in psychology and education. The economists belonged to the Association for Public Policy Analysis and Management! (Cook T.D. (1997), *Lessons Learned in Evaluation over the past 25 Years*, Sage Publications.



In evaluation for effectiveness the evaluator is faced with answering the question whether a particular intervention *caused* a particular result, or put another way, whether a change observed is attributable to the intervention. This kind of cause-and-effect question usually calls for methods that allow findings or estimates to be linked to interventions as closely and conclusively as possible. On the other hand, purposes such as strengthening institutions, improving agency performance, or helping managers think through their planning, evaluation and reporting tasks call for evaluation methods that will improve *capacity* for better performance. Knowledge seeking evaluations involve gaining greater understanding of the issues confronting public policy. The effort to gain such explanatory insights requires strong designs and methods, usually involving both quantitative and qualitative approaches, and advanced levels of both substantive and methodological expertise<sup>30</sup>

This discussion shows that departments with different missions and outcomes are likely to need differing mixes of expertise to carry out their evaluation functions. Departments will need to draw on different disciplines in the establishment of units matched to their particular outcomes. This will need to carry over into recruitment and in-house training programs. Further training in policy analysis would be required.

In the social policy area MSD clearly see a partnership with the universities in order to access the necessary knowledge.

### **Summary/conclusions**

1. While other countries have introduced mandatory schemes for monitoring and evaluation in different parts of Government, New Zealand has not done so. In the past, evaluation of policy in an *ex post* sense has been optional for departments. With the passing of the State Sector Act and the Public Finance Act in 1989, the first step was taken to define the responsibilities for outputs. The lack of mechanisms that relate outputs to outcomes has demonstrated the need for more emphasis on basic research and evaluation into outcomes.

2. Departments have been asked to supply more information on an outcome basis. Outcomes are more difficult to measure than outputs. Recent work on social policy is a start to a more outcome-oriented basis for cross-sectoral policy evaluation and wider consultation on research on policy objectives. More development work needs to be done. An interface needs to be established between the researchers and the policy analysts. Policy issues need to be anticipated in advance. Negative incentives need to be removed from the system.

3. The question of mandatory monitoring and evaluation is difficult to resolve. There will be no evaluation without appropriate monitoring. Audit has proposed a voluntary system of performance reporting for a wide range of government 'entities' on the basis of impact evaluation. This seems to be highly desirable. Departments would find it very difficult to meet a central system of *standards* for reporting in the short term. Different departments have different needs as regards disciplinary background and subject matter. Perhaps a trial period is needed.

---

<sup>30</sup> Chelimsky E. (1997), The coming transformations in evaluation, in *Evaluation for the 21<sup>st</sup> Century*, Sage Publications.

4. Priorities over staffing such initiatives would have to be established in departments. The interface between researchers and policy analysts would have to be clarified. In my experience, an agenda needs to be established and cleared at higher levels before topics are followed up and resources deployed. CEOs need to be involved<sup>31</sup>

6. My conclusion is that a government decision is required to at least require all departments and CEOs to prepare for systematic monitoring and evaluation functions at appropriate levels and consider the staffing and training implications<sup>32</sup>. (4398 words)

---

<sup>31</sup> Scott (2001), p.359, says it is not possible for the government to achieve a high degree of strategic coherence as a whole if any of its key constituent organisations are struggling to produce high-quality strategic thinking and management in their individual areas.

<sup>32</sup> Scott (2001), p.350, suggests listing of major evaluations of effectiveness being undertaken by departments (in terms of strategic objectives) ...to allow an assessment of the scope of the evaluation work.