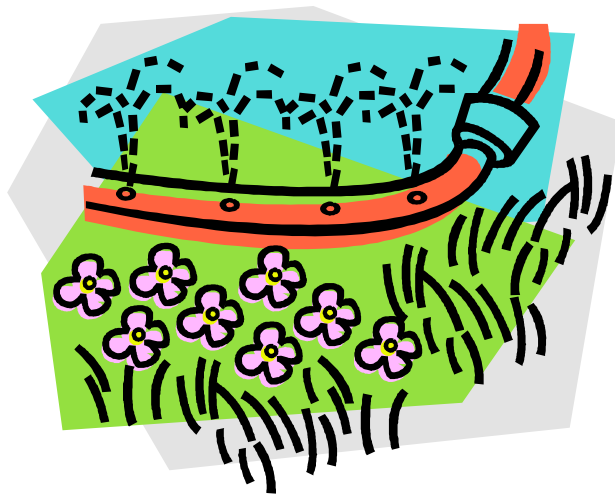


## “The Worth of a Garden”

### Performance Measurement and Policy Advice in the Public Service

[A Discussion Paper]



---

*Prepared by:*

Mark Schacter  
March 7, 2006

*Prepared for:*

Results Management and Reporting Directorate  
Treasury Board Secretariat

# Table of Contents

<b>A. INTRODUCTION</b>	<b>1</b>
<b>B. “HOUSTON, WE HAVE A PROBLEM.”</b>	<b>1</b>
<b>C. NOT A UNIQUE PROBLEM, BUT A DIFFICULT ONE</b>	<b>2</b>
<b>D. WHAT ARE WE TALKING ABOUT?</b>	<b>3</b>
“Policy”	4
“Policy Advisers”	4
“Policy Advice”	5
<b>E. MEASUREMENT OPTIONS – WHAT? &amp; HOW?</b>	<b>5</b>
Performance Measurement Option 1 – Process and Outputs	6
Performance Measurement Option 2 – Outcomes	9
<b>F. EVALUATION VS. PERFORMANCE MEASUREMENT</b>	<b>10</b>
Bibliography	13

# “The Worth of a Garden”. Performance Measurement and Policy Advice in the Public Service

---

*A hundred objective measurements didn't sum the worth of a garden; only the delight of its users did that. Only the use made it mean something.<sup>1</sup>*

## **A. Introduction**

The purpose of this paper is to stimulate discussion on the question of measuring the performance of policy-advice units (“policy shops”) in the federal government. It is based on a review of the literature, including accounts of the experience of other governments (notably Australia and New Zealand) with measuring the quality of policy advice, as well as on the author’s own experience of working on performance measurement and results-based management with a wide variety of Departments and Agencies in the Government of Canada.

The paper lays out the rationale for measuring the performance of policy shops, describes key challenges to measurement in this area, and provides a working definition of “policy” and “policy advice” for the purposes of the paper. The paper then presents options for measuring the performance of policy shops, addresses the strengths and weaknesses of the options, and implementation challenges related to them.

The paper concludes by posing issues for discussion.

## **B. “Houston, we have a problem.”<sup>2</sup>**

After 10 years of ever more rigorous implementation of “results-based management” (RBM) in the Canadian federal public sector, public servants are well used to measuring the performance of programs against their contribution to social/economic results (or “outcomes”) that affect the lives of Canadians. While the process is far from perfect – many weaknesses remain in the implementation of RBM – it has nevertheless become an established part of the program-management routine.

But if your business is the production of policy rather than program outputs, you and your colleagues have probably not felt the RBM imperative to any significant degree. The generally accepted (though not formally stated) understanding is that, as far as performance measurement is concerned, policy shops don’t play by the same rules as program shops.

Why is this so? Two kinds of arguments are typically offered. The first is that policy advice is an art form. Policy advising involves a high degree of judgment, intuition,

---

<sup>1</sup> Bujold (1999).

<sup>2</sup> Apollo 13 astronaut Jack Swigert to Mission Control in Houston on April 14, 1970. The actual words were “Houston, we’ve had a problem here.”

political sensitivity and other ingredients for which there are no objective scales of measurement. "Policy-beauty is in the eye of the beholder", and therefore, the argument goes, judgments about the performance of policy units will be so subjective as to be meaningless for practical purposes.

The second type of argument relates to the "distance" between the outputs of the policy advising process (be they briefing notes, draft policy proposals, memoranda to cabinet, "decks" for Ministers or senior officials, verbal briefings, etc.) and the intended outcomes of the advice (see Figure 1).

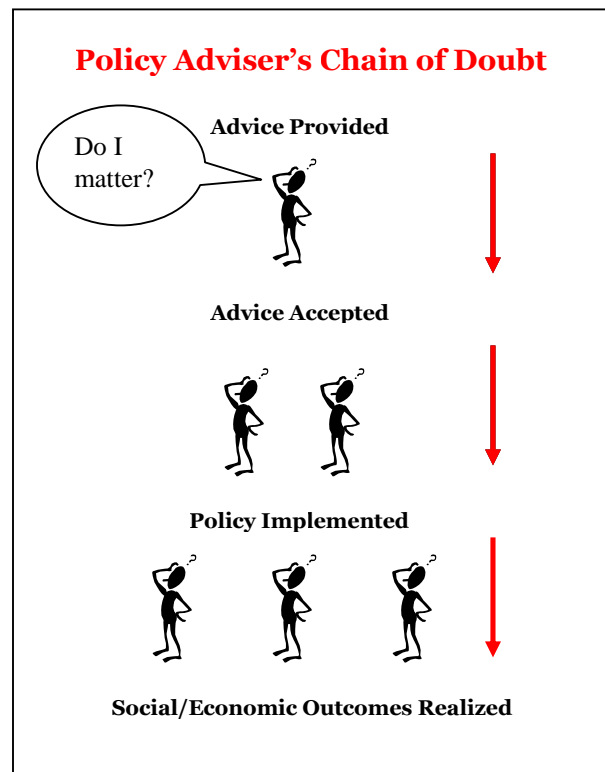
Whether or not policy advice is *accepted* depends on far more than the quality of a given piece (or pieces) of advice; the successful *implementation* of a policy is even more remote from the advisory output; and the *realization of the intended social or economic outcomes* of a policy is still further removed from the advice. A variety of factors – most notably, political ones – that even a highly talented policy advisor may not be able to predict or model may derail the train of events envisioned by the adviser. This gives rise to what is commonly known as a problem of attribution.

Each one of these problems is important in its own right. The combination of the two is doubly intimidating. The case against the feasibility of measuring the performance of a policy unit amounts to this: you have a process that produces outputs for which there appear to be no objective measures of quality, *and*, even if you *could* devise a credible output measurement framework, you would still have to contend with the fact that policy advice is just one strand in an elaborate web of factors affecting policy outcomes. Subjectivity and attribution problems, the two issues that make performance measurement difficult for most forms of public sector activity, are especially potent in relation to measuring the performance of policy units.

### C. Not a Unique Problem, But a Difficult One

It's not surprising, then, that policy shops have escaped much of the pressure faced by program shops to measure and account for their performance against outcomes, as opposed to simply reporting on the use of inputs and the production of outputs. Today, little or no outcome-based performance reporting is being done by policy units in the federal government.

Figure 1



Does this situation make sense? Should program shops have their feet held to the fire on outcomes, while policy shops need only account for their budgets?

On the face of it, there's something unsatisfying about throwing up one's hands in the face of this problem and saying "can't be done." First, policy shops ought to be – and presumably *are* – as focused on outcomes as any program unit, and we ought to be as interested to know about their contribution to outcomes as we are interested in the contribution of program shops.

Second, in an environment where citizens and politicians are increasingly focused on results (as well as inputs and outputs), and are looking for ways to contain or reduce public spending, it is important to be able to advance a strong case for the value-added to Canadian society of policy units in the public sector. Results reporting, for all of its imperfections, is the best available means for doing this.

Third, we know that neither subjectivity nor attribution are necessarily insurmountable obstacles to performance measurement because public servants already deal with them when developing and implementing measurement frameworks for programs. While problems of subjectivity don't normally hinder measurement of program outcomes to the same extent that they would affect the measurement of policy advice, the attribution problem poses as significant a challenge for performance measurement of programs as it does for policy advice.

In short, there is no logical reason to shield policy shops from demands to measure and report on their performance in relation to outcomes. The attendant problems – subjectivity and attribution – are not unique to policy advice; the same issues must be confronted when measuring the performance of programs. While the degree of difficulty may be greater for policy units, this just means that the relevant issue is one of practice rather than principle. The question is not whether, in principle, the performance of policy units can and should be measured (the answer is "yes" on both counts). The question, rather, is *how* one might assess the performance of policy units in a way that makes sense, as a practical proposition, given the obstacles to be overcome and the expected benefits to be gained from the exercise.

## **D. What Are We Talking About?**

Before going further with this exploration of performance measurement and policy advice, it is important to define terms. What is meant by "policy" and "policy advice"<sup>3</sup> for the purposes of this discussion paper and what is the function of a "policy adviser"? What is the basic unit of analysis for the purposes of measurement – in other words, what are the "policy outputs" that will form the basis for judgments about the performance of policy shops?

---

<sup>3</sup> A useful overview is found in Torjman (2005).

## **“Policy”**

There is no universally accepted definition of “policy”. For the purposes of this paper, “policy” is considered to be a set of principles, or the articulation of a high-level course of action, that is intended to influence and determine the general direction of decisions and actions taken by federal Departments and Agencies.<sup>4</sup> A policy provides guidance, or directives, to the executive arm of government on how to address particular aspects of the public interest. So for example, a public policy on reducing Canada’s impact on global warming would be expected to provide Departments and Agencies with guidance on the development of subsidiary policies, programs and regulations related to, say, the reduction of greenhouse gas emissions.

While governments, private companies and not-for-profit organizations all develop and implement policies, this paper concerns itself strictly with “public policies”, that is, policies created and implemented by the government to address matters of public concern. And it is also necessary to recognize the distinction between substantive and administrative policies. Substantive policies are linked to ends, that is, to social and economic outcomes that affect the lives of Canadians. A policy on reducing Canada’s impact on global warming is an example of substantive policy. Administrative policies are linked to management issues, that is to the manner in which the government uses financial resources, procures assets and materiel, hires, rewards and disciplines staff, etc. An example of an administrative policy would be the federal government Travel Policy or the Internal Disclosure Policy. This discussion paper is concerned with substantive policies only.

## **“Policy Advisers”**

Policy advisers are the people who are called upon to support the development of policy. They provide advice – ultimately intended for Ministers – on the nature of public policy problems, the formulation of policies for addressing matters of public concern, and the relative merits of different policy approaches.<sup>5</sup> Policy advisers may be public servants working in Departments and Agencies, political advisers working in Ministers’ offices, or individuals outside of government working in think tanks, consulting firms or universities. This paper focuses on policy advice provided from within the public service (though, presumably, a similar analysis would apply to policy advice from the other two sources).

Tasks undertaken by policy advisers include:

---

<sup>4</sup> It is interesting to note that the literal translation of the Chinese character for policy is “needle of direction.” Briggs (2005).

<sup>5</sup> One policy adviser suggested to me that this was an idealized view of what policy officers do in the Canadian federal government. She said that only a relatively small amount of a policy adviser’s time is devoted to “true” policy work as described here. She said that the majority of an adviser’s time is spent on less substantive tasks such as responding to last-minute Ministerial requests for information, dealing with correspondence, etc. Be that as it may, this discussion paper focuses on the “core” policy-oriented functions of the policy adviser.

- definition (or re-definition) of policy problems;
- analysis of policy problems; taking problems apart and “reassembling” them in ways that facilitate understanding and action on the part of decision-makers;
- developing options for dealing with an issue on which the government must develop a view and initiate action;
- communicating the results of policy analysis in a way that facilitates decision-making.

### **“Policy Advice”**

The performance of a policy shop, and of the policy staff who work in one, will be judged against the quality of the policy advice that they provide. So “policy advice” is the basic unit of account for the purposes of performance measurement. If that is the case, then what “counts” as a unit of policy analysis? What kinds of outputs will be the basis against which to measure performance: memoranda to cabinet? draft policy documents? written briefings? oral briefings?

This paper will not attempt to address this question. It will work from the assumption that “we know policy output when we see it.” Opportunities will be sought during the discussion prompted by this paper to elicit views on the feasibility of and/or value to be added by developing a more precise definition of “policy advice”.

### **E. Measurement Options – What? & How?**

As with any production activity, the process of creating policy advice can be broken down into four elements: consumption of inputs; transformation of inputs into outputs; quality of outputs; and contribution of the outputs to outcomes (or “results”). Each element provides a basis for measurement. One could thus choose to measure any one of, or any combination of:

- a policy shop’s use of financial and human resources (does it operate on budget? how efficiently does it produce outputs?);
- the process by which policy outputs are created;
- the characteristics of the policy outputs in relation to quality standards, or, in the absence of standards, to generally accepted views about quality;
- the degree to which policy outputs are believed to have contributed to intended policy/program outcomes.

The first element is routinely integrated into the management and performance measurement of all public sector activities, and is not of interest for the purposes of this paper. The discussion here will focus on the latter three elements.

The analysis of how best to incorporate these elements into measuring the performance of a policy shop is tied to the two key obstacles, noted above, that affect performance measurement: subjectivity and attribution.

Of the two obstacles, attribution is the fundamental one. The attempt to link outputs to outcomes – directly or indirectly – is fundamental to all performance measurement, whether of program or policy shops. The ultimate purpose of all public activity is to have a positive impact on Canadian society. The ultimate test of performance, therefore, is evidence – direct or indirect – of a link between the program or policy output in question and the social or economic outcomes to which it is supposed to contribute. So the challenge of attribution, though daunting, is inescapable if performance measurement is to be credible and meaningful.

As Figure 1 suggests, the attribution problem impedes measurement of the performance of policy shops in relation to evidence of a direct contribution to social or economic outcomes. So many other actors and factors are at play, it is argued, that isolation of the impact of a particular piece of advice, and production of direct evidence of a link with outcomes, are practically impossible. If one accepts this proposition, then the alternative is to focus on the other two elements – process and output quality – and then make assumptions about the link between process/output on the one hand, and outcomes on the other.

The outcome option should not, however, be dismissed too readily. Even if it is not feasible to measure performance in relation to ultimate outcomes, it may be possible to go beyond process and outputs, by measuring performance in relation to immediate and intermediate outcomes. Then, as with process/output measurement, assumptions would have to be made about the link between (in this case) lower-level outcomes and ultimate outcomes.

So let's now consider, in more detail, the two primary options for measuring the performance of policy shops: the first option focuses on process and outputs; the second option includes process and outputs, but also incorporates outcomes.

### **Performance Measurement Option 1 – Process and Outputs**

This option addresses the subjectivity problem by proposing a relatively objective set of criteria for assessing the quality of policy advice. It addresses the attribution problem by proposing to use the quality of policy advice as a proxy for contribution to social and economic outcomes. (In other words, an assumption is made that if policy advice is “good” in relation to an agreed set of criteria, then it is reasonable to conclude that policy advice is leading to good policy that leads, in turn, to desired outcomes.)



This option depends upon the premise that it is possible to reach some minimum degree of consensus on key criteria of a "good" process for developing policy advice, key criteria of "good" policy-advice outputs, *and* that these criteria will be applied in a similar fashion by different observers.

We know that creators and consumers of policy outputs have a tacit understanding of what constitutes "good" process and product. When users of policy advice praise a piece of work, or criticize it, they do so on the basis of some notion of what "good" looks like. Similarly, producers of advice use their own notion of "good" as a basis for guiding their approach to the task. While there will be some element of idiosyncrasy – two policy practitioners or consumers will not have identical views on the criteria for "good" work – interaction among policy practitioners, among policy consumers, and between practitioners and consumers has nevertheless given rise to a core of shared ideas about "good" process and output. While we may lack definitive quality standards for policy advice, the common ground is significant enough to allow us to assume a broad consensus.

One view about process and output standards for policy advice on which there is some agreement in the policy community<sup>6</sup> holds that good policy advice:

- is **timely** – it is ready when Ministers and other decision-makers need it;
- has been prepared based on adequate **consultation** with stakeholders inside and outside government;
- clearly articulates the **purpose** for which it has been prepared;
- has a sound **logical basis** – there is a clear description of the facts and assumptions upon which the advice is based, and a clear articulation of the links between fact and assumptions on the one hand, and conclusions and recommendations on the other;
- has a sound **evidentiary basis** – the underlying evidence is accurate and complete;
- is **balanced** – it presents a representative range of viewpoints;
- presents an adequate range of **viable options for action**;
- is **relevant** to the current situation faced by decision-makers – it takes into account current realities (including political realities) and anticipates related developments;
- is **well presented** to the reader – the prose is concise; the text is well organized;

---

<sup>6</sup> This is adapted from informal standards that have been proposed in various documents (see Bibliography) related especially to efforts by the governments of New Zealand and Australia to measure the quality of policy advice.

- it **pragmatic** – it bears in mind implementation issues.

If we allow that this reflects consensus on a basis for measuring the quality of policy outputs, then the question becomes *how* to measure. Who is best qualified to make judgments relative to these criteria? From among the variety of opinions that might be sought, whose opinions matter most? (This is where the assumption that quality criteria will be applied in a similar fashion by different observers begins to break down.)

It would be natural to suppose that the most important (even if not necessarily the most knowledgeable) opinion about the quality of policy advice would belong to the “client” of the policy shop. Who is the client? The obvious answer is the Minister, inasmuch as he is the ultimate intended recipient of a policy shop’s output. But there are many layers between the policy shop and the Minister, so one would also have to give weight to the judgments of senior officials.

It might also be argued that there are many “indirect” clients of policy advice. These would be stakeholders – people inside or outside the government who would be directly or indirectly affected if a given piece of policy advice were to be accepted and implemented. The extent to which they should be involved in judgments about the quality of policy advice is an interesting question.

As well, there are parties other than clients and stakeholders who may have a role to play in assessing the quality of policy advice. These would be technical experts – people such as academics, Departmental specialists and policy analysts working in non-governmental organizations – who have specialized knowledge of the subject matter in question.

***The strengths of this option for measuring the performance of a policy shop are that it:***

- *focuses on things – processes or outputs – that are largely within the control of the policy shop;*
- *is founded on consensus about the key ingredients of good policy advice; and*
- *is built on a reasonable assumption that “good” policy advice is likely to make a significant contribution to social and economic outcomes.*

***The weaknesses of this option are that it:***

- *focuses on means (outputs) rather than ends (outcomes), and relies on an assumed link between outputs and outcomes;*
- *it cannot entirely overcome the problem of subjectivity, and so leads to difficult questions about how to weight the different opinions of different assessors of policy advice;*

- *leads to difficult questions about how to weight the different components of policy-advice quality (some are undoubtedly more important than others; and the relative importance of the components may vary according to circumstances); and*
- *will be time-consuming and costly to implement on a regular basis, because it depends upon seeking the views of individuals.*

### **Performance Measurement Option 2 – Outcomes**

This option deals with both the subjectivity problem and the attribution problem by going directly to the issue of outcome “ends” as opposed to process/outcome “means”. It focuses on a (relatively) objective assessment of whether or not outcomes have occurred, thereby avoiding the subjective question of whether or not policy advice is “good”. It addresses the attribution problem by proposing to use immediate and intermediate outcomes as proxies for contribution to social and economic outcomes. There are two assumptions implicit in this option: first, that an immediate outcome (e.g. a decision by the Minister to adopt a particular policy position) was caused to a significant degree by a given policy output; and, second, that an ultimate social or economic outcome will be positively influenced in some way by the Minister’s decision.

This option emphasizes the steps between the production of policy outputs and the occurrence of social or economic outcomes that affect the lives of Canadians. The intermediate steps would be ones such as the acceptance of policy advice by decision-makers (or acknowledgement by decision-makers that the policy advice was an aid to decision-making) and the implementation of the related policy. The rationale for focusing on these low-level outcomes would be that the policy shop is presumed to have some reasonable degree of direct influence over them (whereas it would be presumed to have no control over ultimate outcomes).

So the key measurement questions, under this option, would be related to the acceptance and implementation of policy advice. A policy shop would be judged to be performing well if its advice was accepted and/or implemented.

***The strengths of this option for measuring the performance of a policy shop are that it:***

- *emphasizes ends rather than means;*
- *seeks to isolate outcomes over which the policy shop may have a reasonable degree of control; and*
- *is founded on a relatively objective assessment of whether outcomes have occurred or not.*

***The weaknesses of this option are that it:***

- *depends heavily upon a sometimes questionable link between policy outputs and outcomes;*
  - *factors other than the quality of the policy advice may have an important impact on whether or not policy advice is accepted and/or implemented;*
- *assumes that it is possible to distinguish between the impact on policy decisions of multiple advice inputs (because any given policy decision may have been influenced by advice from sources other than the policy shop whose performance is being measured); and*
- *assumes that it is possible to distinguish between the impact of "old" and "new" policy advice (for example, a policy decision might be affected by advice given years before).*

**F. Evaluation vs. Performance Measurement**

As the preceding analysis suggests, there are major challenges associated with using performance measurement as a tool for assessing the value added to society by policy advice. The most daunting challenge relates to attribution. While this issue also affects performance measurement of programs, the challenge is especially acute with regard to policy advice (because policy advice is one step further removed from outcomes than program outputs). Attributing the actions of a decision-maker, especially a Minister who must strike a balance between technical and overtly political considerations, to a given piece or source of policy advice is a difficult proposition at the best of times. Attempting to go further and attribute a decision-maker's action to the *quality* of a piece of policy advice – assuming that consensus can be reached on the defining characteristics of "quality", how to apply them to a given piece of policy advice under a given set of circumstances, and on who is the best judge of quality – is more difficult still. And the degree of difficulty becomes even higher when an attempt is made to link social/economic outcomes with advice outputs.

The key factor distinguishing a policy shop from a program shop for the purposes of performance measurement is that the former's work is characterized by "low specificity"<sup>7</sup> – i.e. it is difficult if not impossible to specify (i) a standard routine by which the work should be done; (ii) a standard set of output-quality criteria; and (iii) a standard and well-defined path, or logic, by which outputs lead to outcomes. Low-specificity organizations and tasks pose especially difficult problems for performance measurement – problems so significant that it may be impractical (if not impossible) to apply standard performance measurement in a way that yields useful results.

---

<sup>7</sup>This concept was first addressed in detail in *Institutional Development. Incentives to Performance*, by Arturo Israel, Baltimore: Johns Hopkins University Press, 1987.

This does not mean that one should not attempt to assess the quality of a policy shop's performance. But it does suggest that evaluation may be worth considering as a better tool than performance measurement for this particular task. Evaluation, though closely related to performance measurement, differs from it in ways that may provide a better fit with the subtleties and ambiguities of the policy-advice process.

Performance measurement and evaluation differ in their level of analysis and their time-frames. Performance measurement describes performance, while evaluation seeks to explain it. Performance measurement focuses on current performance and produces regular reports on an annual basis (if not more frequently), while evaluation looks back at past performance, and evaluation reports are produced over longer intervals in order to allow for a broader and richer evidence base with regard to outcomes. Performance measurement is speculative about the contribution of outputs to outcomes, while evaluation aims to be definitive.

While performance measurement relies on the repeated collection of a uniform set of data in relation to agreed-upon performance indicators, evaluation probes more deeply while also being more rigorous and more open-ended, using multiple lines of evidence to reveal causal links that provide a basis for credible claims of attribution.

The basic performance measurement questions in relation to a policy shop would be: How well is the policy shop performing now, in relation to intended social and economic outcomes? Based on what we know now, do the its outputs appear to be contributing to outcomes?

By contrast, the basic evaluation questions for a policy shop would be: Given the body of accumulated information – from a variety of sources and perspectives – about the performance of the policy shop over a period of years, would it be reasonable to conclude that its outputs have in fact made a significant contribution to achieving intended outcomes? If they did (or didn't), then why was the policy shop successful (or not)?

Evaluation is more costly and time-consuming than performance measurement. The trade-off is that it is done periodically – whereas performance measurement is done on an ongoing basis – and provides a much more reliable basis for assessing the performance of a highly non-specifiable activity such as policy advising.

## **G. Questions for Discussion**

- What types of outputs should be counted as “policy advice”?
- Does the set of process and output characteristics of “good” policy advice proposed on p. 7 provide a reasonable basis for assessing the quality of policy advice?
  - Are there elements that should be removed? added?

- Are some of the ingredients of good policy advice more important than others? To what extent will this depend upon circumstances?
- Whose opinions about the quality of policy advice should be considered? Whose opinions should carry the greatest weight?
- Is it preferable to focus on measuring performance against the immediate and intermediate outcomes of the policy-advising process, or is it better to give greater emphasis to process/output quality standards?
  - Do policy shops have enough control over immediate/intermediate outcomes (acceptance of advice; implementation of advice) to allow for their use as a basis for performance measurement?
- Given the major challenges to continuously measuring the performance of policy shops through standard performance measurement techniques, would it make more sense to focus on periodic evaluation of selected policy shops?
- All things considered – taking into account the costs and special challenges of measuring the performance of policy shops, and weighing these against the anticipated public benefits of measuring policy shops' performance – does it make sense to impose performance measurement or evaluation on policy shops, or should they be exempted from any requirement to measure or evaluate their performance?

## **Bibliography**

Australian National Audit Office (2001). "Some Better Practice Principles for Developing Policy Advice," Canberra.

\_\_\_\_\_ (undated). "Performance information for policy advising functions," Canberra.  
[www.anao.gov.au](http://www.anao.gov.au)

Behm, Allan, et. al. (2000). "A value-creating model for effective policy services," *Journal of Management Development*, Vol. 19, No. 3

Briggs, Lynelle (2005). "A Passion for Policy," (Australia and New Zealand School of Government lecture series), [www.apsc.gov.au/media/briggs290605.htm](http://www.apsc.gov.au/media/briggs290605.htm)

Bujold, Lois McMaster (1999). *A Civil Campaign*, Riverdale (NY): Baen Publishing.

Bushnell, Peter (1998). "Does Evaluation of Policies Matter?" (Address to Policy Evaluation Workshop at the Foreign and Commonwealth Office (London) )

Davenport, Thomas H. (2005). *Thinking for a Living*, Boston: Harvard Business School Press.

Department of Finance and Administration (Australia) (undated). "The case of policy advice." [www.finance.gov.au](http://www.finance.gov.au)

Department of Treasury and Finance (Tasmania) (1997). "Guidelines for the Definition, Costing and Reporting of Policy Advice Outputs," Hobart.

Di Francesco, Michael (2000). "An Evaluation Crucible: Evaluating Policy Advice in Australian Central Agencies," *Australian Journal of Public Administration*, Vol. 59, No. 1.

\_\_\_\_\_ (2000a). "Performance Management and Policy Advice: Learning for Policy Coherence?" (Paper prepared for the third bi-annual research conference of the International Public Management Network.)

\_\_\_\_\_ (1999). "Measuring performance in policy advice output. Australian developments," *The International Journal of Public Sector Management*, Vol. 12, No. 5

Hawke, G.R. (1993). *Improving Policy Advice*, Wellington: Institute of Policy Studies.

Israel, Arturo (1987). *Institutional Development. Incentives to Performance*, Baltimore: Johns Hopkins University Press.

Johnson, Robin (2003). "The Evaluation of Economic Policy." (Paper prepared for Annual Conference of the New Zealand Association of Economists.)

Nagel, Stuart (2001). "Conceptual Theory and Policy Evaluation," *Public Administration and Management*, Vol. 6, No. 3.

Office of the Provincial Auditor (Manitoba) (2003). "A Guide to Policy Development," Winnipeg.

\_\_\_\_\_ (2001). "A Review of the Policy Development Capacity Within Government Departments," Winnipeg.

Scott, Claudia (2005). "Value-Adding Policy Analysis and Advice: New Roles and Skills for the Public Sector." (Paper prepared for Association for Public Policy Analysis and Management annual research conference.)

State Services Commission (New Zealand) (2000). "Building Advice: The Craft of the Policy Professional," (Working Paper No. 7), Wellington.

\_\_\_\_\_ (1999). "Essential Ingredients: Improving the Quality of Policy Advice," Wellington

\_\_\_\_\_ (1992). "The Policy Advice Initiative: Opportunities for Management," Wellington.

Torjman, Sherri (2005). "What is Policy?" Ottawa: The Caledon Institute of Social Policy.

Treasury Board Secretariat (2006). "Performance Measurement Tools for Government of Canada Communications: Procedures Manual" (draft), Ottawa.

Uhr, James and Keith Mackay (eds.) (1996). *Evaluating Policy Advice: Learning from Commonwealth Experience*, Canberra: Federalism Research Centre, Australian National University.