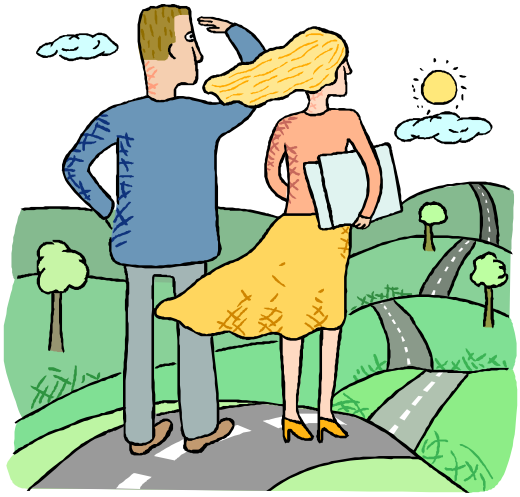
What Will Be, Will Be

The Challenge of Applying Results-based Thinking to Policy



By Mark Schacter

Institute On Governance
Ottawa, Canada

www.iog.ca

The views expressed here are the responsibility of the author and do not necessarily reflect those of the Institute On Governance or its Board of Directors.

What Will Be, Will Be
The Challenge of Applying Results-based Thinking to Policy

ISBN 1-894443-10-1

Published and distributed by:

The Institute On Governance
Ottawa, Ontario, Canada

Phone: (1-613) 562-0090
Fax: (1-613) 562-0097
Web Site: www.iog.ca

Table of Contents

Part 1 – Why It Matters

When I was just a little girl,
I asked my mother, "What will I be?
Will I be pretty?
Will I be rich?"
Here's what she said to me:

"Que sera, sera,
Whatever will be, will be;
The future's not ours to see.
Que sera, sera,
What will be, will be."[1]

A. Introduction

Public policy practitioners have a lot in common with both the little girl and the mother in the song.

Like the girl, public policy practitioners are obsessed with the future. The practice of policy is all about tomorrow. It is about seizing opportunities, anticipating problems, and, ultimately, improving the lives of citizens. Policy practitioners devote their professional lives to trying to make the future into something better than the present. "What will be?" is a question that keeps them awake at night.

Like the mother, policy practitioners wish they could see the future, but know that they can't. There is much they want to affect, and little they can control. A central challenge for policy practitioners, as for parents, is to manage the tension between what they can influence and what they can't; between "what will be" (no matter what) and "what might be" (if they do their job well). (The parallels between parenting and policy-making are so strong that I will draw upon them in detail later in the paper.)

If you were to ask a policy practitioner, "How is your work going?", you might find that he or she pauses before answering. The question you are posing is about today, but the policy practitioner lives in the future. "Ask me in five or ten years," might be the reply.

"How is your work going?" is a question about *performance measurement*. Implicit in it is a request for information about how well you, your colleagues, your work unit or your organization are performing *now*. It is a natural and reasonable question to ask, and yet policy practitioners often feel disinclined or ill-equipped to respond.

[1] From the song "Que Sera Sera," by Jay Livingstone and Ray Evans. It was made famous by Doris Day in the 1956 Alfred Hitchcock film *The Man Who Knew Too Much.*

This paper will argue that policy practitioners' reticence about performance measurement, though understandable, is unnecessary and unjustified. Having made that point, the paper will then try to *help* policy practitioners think their way through the challenges of applying results-based management (RBM) and performance measurement concepts to policy. The paper will describe the challenges in some detail, and then propose practical approaches to overcoming them.

B. Policy Practitioners Under Pressure

> Canadians have a right to know *what* governments are trying to achieve, *why* governments believe certain activities contribute to their objectives, and *how* governments plan to measure whether they are achieving the objectives.[2]

The quotation is a recent official statement of the Canadian government's position on results reporting in the public sector. It follows from a broader management philosophy regarding "results-based management" (RBM) which the government began promoting aggressively in the mid-1990s, and which is now conventional wisdom (at least at the rhetorical level).[3]

Although RBM is on its way to becoming embedded in the management culture of the Canadian public service, a significant gap remains. Public sector managers remain less comfortable applying the principles of RBM and performance measurement to *policies* than to *programs*. The divergence in attitudes among program and policy practitioners regarding RBM and results-reporting is as curious as it is important. Programs normally follow from policies. It is common to talk about the "*policy* basis" or the "*policy* objectives" of a program. Given our assumptions about the tight logical connections between program and policy, why should it seem more natural to apply RBM and performance measurement to the one than to the other?

Against this background, this paper aims to respond to a growing need among public sector policy managers. We see signs of policy practitioners feeling under increasing pressure to "demonstrate a results orientation" and to report on the "results" of their work.[4] But policy practitioners who seek guidance on this from the literature or from the current body of professional "best practice" report that they come up empty handed. The material they find is inevitably couched in terms of programs rather than policies.

This paper is written with the aim of helping this pressured and perplexed group of policy practitioners think their way through the problem of applying RBM and performance measurement to policy. It is targeted to the managers and staff of "policy units" within the government. The responsibility of these policy units varies across Departments, but typically includes the development of policies (or policy advice and guidance) in support of the overall Departmental mission or of particular Departmental programs.

---

[2] *Canada's Performance 2001,* (President of the Treasury Board, Annual Report to Parliament), Ottawa: Treasury Board Secretariat, 2001.
[3] See *Results for Canadians,* Ottawa: Treasury Board Secretariat, 2000.
[4] For example, in the workshops that the Institute On Governance runs for public servants on RBM and performance measurement, there has been a noticeable increase in demand from participants for sessions that focus specifically on the question of measuring the performance of policy units.

In the spirit of being pragmatic, I make no attempt in this paper to be "rigorous" in the academic sense. The paper offers insights based on actual experiences of working with public servants on performance measurement problems, and on my own sense of what is feasible at the operational level. It makes no reference to the literature from the related field of policy evaluation. Given the cries for support coming from the federal-government policy community, one can only assume that the policy evaluation literature – whatever its merits – is not seen by practitioners as helping them wrestle with the immediate problem of applying RBM and performance measurement to policy[5].

The conceptual basis for this paper is found in a companion piece entitled "Not a Tool Kit. Practitioner's Guide to Measuring the Performance of Public Programs."[6] Readers may want to browse through "Not a Tool Kit" before reading this paper. Rather than restating in full the ideas contained in that document, I will only make brief reference to it where necessary.

C. What's the Problem?

Why is there discomfort related to applying RBM and performance measurement to policies? What creates an assumption in the minds of some policy practitioners that the principles of RBM and performance measurement, though applicable to programs, are not a "good fit" with policy work?

Without conducting a representative survey of policy practitioners, and in the absence of a useful body of literature on this question, I cannot pretend to offer anything approaching an authoritative answer. I can however surmise, based on interactions and conversations with policy and program practitioners over the years.

Policy practitioners who are hesitant to apply performance measurement to their work will sometimes argue that policy work is so unique – so different from program work – that it is not reasonable to subject it to the same rules of RBM. The uniqueness argument is offered in many guises, but normally boils down to some or all of the following three premises:

- policy is intangible;
- policy-making is highly subjective;
- understanding the impact of policies is a complicated and messy business.

The related arguments are typically along the following lines:

Policy is Intangible. The Oxford English Dictionary defines "policy" as "a course or principle of action adopted or proposed by a government, party, business or individual." You cannot see or

[5] Our own review of the literature revealed a dearth of writing on the general question of RBM and performance measurement as they apply to policy.
[6] "Not a Tool Kit. Practitioner's Guide to Measuring the Performance of Public Programs," by Mark Schacter, Ottawa: Institute On Governance, 2002. Available at www.iog.ca/publications/guide.pdf

touch a "course or principle of action" (except on paper!). You cannot watch it working in the way that you can observe the activities and effects of public programs. You therefore cannot measure its performance in any useful or meaningful way.

Policy-making is Subjective. Good policy-making demands large inputs of human intellect and analytical power. As well, although policy-making is almost always informed by data, it also (and unavoidably) is built upon a framework of opinion and judgement. The inherent subjectivity of policy-making means that the performance of policies – unlike the performance of programs – cannot be measured in any systematic way.

Understanding the Impact of Policies is a Complicated and Messy Business. What will be, will be. The path from formal approval of a policy to the realization of the policy's objectives will often be long, indirect and uncertain. Many factors – some known and predictable, others unknown and unpredictable – will affect the attainment of targeted objectives. Years may pass before policy-makers get a sense of whether the results they had hoped for are materializing. And even then, it may be difficult to say whether the results were caused by the policy itself, or by other unrelated factors. Under such circumstances, one would have to be either exceptionally brave or uncommonly foolish to attempt to measure the performance of policies.

D. Yes, But . . .

Intangibility, subjectivity and uncertain causation are indeed touchstones of the policy business. They help us understand why it may often be difficult to apply the concepts of RBM and performance measurement to policy. But they are *not*, as some would argue, reasons for avoiding RBM and performance measurement. The argument that it is practically impossible, from a technical perspective, to apply performance measurement to policy does not hold water.

Factors such as intangibility, subjectivity and tangled causation are not unique to policy. They are often also a feature of programs. While some social and economic programs funded by the government lend themselves easily to measurement and causal attribution, many (perhaps most) do not.[7] The challenges inherent in public-sector performance measurement are not therefore peculiar to policy, but apply to programs as well. For the technical purposes of performance measurement, the differences between policies and programs are matters of degree rather than kind. So if we agree that RBM and performance measurement can be made to work for programs, then, *ipso facto,* we must agree that they can be made to work for policy.

There is also a powerful *governance argument* for imposing RBM and performance measurement on policy with the same rigor that we insist on applying it to programs. Development of both policies and programs is made possible by public funding; both are produced by the government with certain objectives in mind. Whenever we have publicly funded activities undertaken with particular purposes in mind, it becomes natural for taxpayers to want to ask: "How are you performing? Are you accomplishing what you set out to accomplish with our money?" To tell the taxpayer that we would be willing to answer this question in relation to programs but not policies would be absurd! Whether we are talking about policies or

[7] To take but one example, consider programs funded by the federal government aimed at supporting Canadian cultural development and national unity

programs, the political accountability dictum cited at the head of Section B remains in effect. Taxpayers fund the time and resources devoted to policy-making, as much as they fund programs. In relation to policies as well as to programs, therefore, taxpayers have a right to expect credible answers to the "what", "why" and "how" questions noted above. Policy practitioners in the public sector are no less obliged to be accountable for their performance than are their program counterparts.

Part 2 – The Major Challenges

A. Parents and Policy-Practitioners: Birds of a Feather

This section focuses on the key challenges you are likely to face as you attempt to apply the principles of performance measurement to policy. It is worth repeating that each one of these challenges might be equally applicable to measuring the performance of a program. Having said that, it is also true that the challenges described here will often tend to be most acute in relation to policy. As well, it is likely that *all* of the challenges discussed in this section are likely to occur *simultaneously* in relation to a policy. (Whereas, in relation to a program, one might have to deal only with one or two.)

I am going to use a familiar metaphor – parenting – as a device for describing and explaining the key challenges related to measuring the performance of policy. This has at least two advantages.

First and most important, it is a robust metaphor for our purposes. It illuminates issues of performance measurement that are strikingly similar to those faced by policy practitioners.

Second, many of us would feel that we have a good common sense understanding of the parenting "business." We can relate to it easily, and can think about it with confidence. Basing our discussion in a subject that appears familiar and intuitive will help to push aside the unjustified veil of mystery that surrounds public policy. It is helpful to dispel the myth – sometimes used as an argument for rejecting the application of performance measurement to policy – that policy-making is such a specialized domain that people outside the discipline cannot understand it.

Let's pursue the parenting metaphor via a thought-experiment. Let's imagine that a thoughtful and objective observer, trained in the art and science of performance measurement and RBM, is assigned to my household for an extended period. His job is to measure my and my wife's performance as parents. How would he carry out his assignment, and what types of challenges is he likely to face? The following sections address these questions. As we review the key challenges – there are four – the parallels to measuring the performance of policy should be self-evident.

B. Challenge No. 1 – Multiple High-level Outcomes

The first order of business for our observer would be to understand the high-level outcomes that guide our parenting activity. Before he can begin to think about measuring our performance, he needs a clear picture of the results we are seeking.[8] What are we trying to accomplish as parents? It is only in relation to those intended outcomes – whatever they may be – that our resident observer can judge our performance. In the absence of clarity around purposes – or "ultimate outcomes" – performance measurement is meaningless.

[8] "Not a Tool Kit", pp. 13-14.

As we raise our children, what are we aiming for? What outcomes are we seeking? Do we want a happy child? a smart child? a sociable child? an athletic child? some combination of all of these?

We will also have goals in mind for our children that extend beyond their childhood to adolescence and adulthood. What kind of adults do we want our children to turn out to be?

Like many parents, we might have a range of high-level outcomes in mind for our children. Perhaps we are hoping that they will turn out to be professionals of some sort. Or perhaps we are hoping they will be successful entrepreneurs. We may want them to marry and raise families of their own, and provide us with grandchildren. We may hope that our children will develop into adults who are well liked and widely regarded as "good people."

The reality of course is that we, like most parents, will be seeking some combination of "high-level outcomes" in relation to their children. A few possibilities have been suggested above, but the list of conceivable desired outcomes is open-ended! And several of the outcomes that we choose to pursue may have equally high priority for us. What our observer may find, therefore, is that there is no one single obvious high-level outcome against which to measure our performance. Any number of different objectives might apply, either *sequentially* or *simultaneously*. In the presence of multiple, high-level outcomes, some of which may be conflicting, performance measurement becomes an extremely difficult matter.[9]

C. Challenge No. 2 – Measuring the "Unmeasurable"

Let's assume that our observer is able to discern a set of high-level outcomes that my wife and I cherish in relation to our children. Suppose he concluded that more than anything else, we wanted our children to be "happy", to be regarded as "good people", and to be "confident and self-assured". He then takes these high-level outcomes as the basis against which to measure our performance.

This is fine as far as it goes, but how does our observer measure our performance against outcomes of this nature? Can performance against qualitative and subjective outcomes such as these be measured in any meaningful or credible way?

Outcomes of this type do indeed pose special measurement challenges. Measurement, an activity we normally associate with objectivity and precision, becomes controversial, ambiguous and open to debate when outcomes include important qualitative and subjective elements.

[9] There is an important issue related to measuring performance against high-level outcomes which I will not address in this paper. It has to do with the *quality* of the outcomes themselves. Consider an extreme and unlikely example: parents who want their child to grow up to be a terrorist. If the child grows up to be a terrorist, should we conclude that the parents performed well, because the high-level outcome was achieved? Or do we say that the parents failed because they chose a bad outcome? Strictly speaking, performance measurement takes high-level outcomes as given, and so in this case we might be constrained to concluding that the parents performed well. If we were to conduct an evaluation, however, we would ask deeper questions about, among other things, the initial choice of ultimate outcomes.

Our resident observer, however, has no choice but to rise to the challenge. The high-level outcomes that we as parents have decided to focus on are valid and important in the context of parenting. It is entirely legitimate to think that they would be significant indicators of our performance as parents. Just because they are "soft" outcomes does not mean that they are necessarily any less important or valid than "hard" outcomes that are more readily quantifiable and therefore more easily measured. (For example, outcomes such as "I want my child to grow up to be rich" or "I want my child to get straight 'A' grades in university" would be much less problematic from a performance measurement perspective. But they are not necessarily any more valid or important than the outcomes we have chosen in this example.)

D. Challenge No. 3 – Time Lag

It takes a while (we assume) for our child-rearing activities to bear fruit. The lessons we teach our children, the knowledge we impart, the affection we show, the praise and punishment we dole out – all of our parenting activity takes time to show results. Sometimes, the results we are seeking may not become apparent for many, many years. Indeed, there may be times when, as parents, we become convinced that our children are not developing in the ways we had hoped for. There may be prolonged periods of time where they appear unhappy, are poorly behaved, seem poorly adjusted, etc.

Ultimately, of course, these negative trends in our children's behavior may be aberrant "phases" that they are passing through. Over time, their development may proceed in the more positive direction that we had been hoping for.

In the language of RBM, there is often a long time-lag between our parenting *activities* of today, and the ultimate *outcomes* that we hope to see in our children many years down the road. For our resident observer, this hard reality of the parenting business poses a large performance-measurement challenge. His job is to measure our performance *now*, even though the results that will be definitive indicators of our performance may not be visible until some distant future time.

E. Challenge No. 4 – Attribution

The attribution problem arises whenever there is a significant gap between *affecting* an outcome and *controlling* an outcome. It is perhaps the most important and difficult challenge to performance measurement[10].

Our children's development will be affected by innumerable factors beyond the range of what we as parents can control or even influence. While we very much want to *affect* the development of our children, it almost always seems to us that there are too few things over which we have *control* with respect to how our children will end up! This is *the* fundamental challenge – and frustration – for parents, much as it is for policy practitioners in the government. In both cases you are deeply engaged in trying to affect outcomes over which you may not, in fact, be able to exert very much control.

[10] See "Not a Tool Kit", pp. 18-19 and 23-26 for a detailed discussion of the attribution problem.

Consider the difficulty that the attribution problem creates for our resident observer. What if our child turns out badly as an adult – a serious criminal, for example? How would he rate our performance as parents. Does the undesirable "result" mean that we have performed poorly?

The answer is not obvious. It is not inconceivable to have a situation where good parental behavior does not produce a "good" child. We might have done everything that any reasonable person might have expected of us as parents: we may have been kind and loving, attentive to our children's intellectual and emotional development, provided them with all of the opportunities that were possible for us to provide, etc. Despite our good efforts, our child may turn out to be the opposite of what we had hoped.

This is the unavoidable reality of activities – such as parenting and policy-making – where outcomes have many causes, and where the chains of causation may be tangled and practically impossible to unravel. In the case of our children, far less than 100 percent of the outcome is going to be attributable to what we have done as a parent. If our children turn out to be wonderful adults, we cannot take all of the credit; if they turn out poorly, it wouldn't be reasonable to assign us all of the blame.

Part 3 – Dealing With the Challenges

A. Breaking Down the Problem

Anyone who has ever worked in a policy unit will have confronted some or all of the four challenges described above.

- It is in the nature of public policies that they will be simultaneously pursuing *multiple, high-level outcomes*, some of which may even be in conflict with each other. There are many ways in which this might arise, but the most troublesome situations occur when there is inconsistency or conflict between the intended political and technical outcomes of a policy. "Not a Tool Kit" gives the example[11] of a public housing policy where the intended and unspoken political outcome (eliminating housing as a contentious election issue) was seen by the Minister as being far more important than the explicit technical outcome (building the targeted number of public housing units).

- It is not uncommon for policies to pursue outcomes that can only be described in qualitative and subjective terms, and that may therefore appear to be *unmeasurable.* For example, the mission of the Department of Canadian Heritage is to contribute to "a more cohesive and creative Canada." One has to assume that all of the policies developed by the Department are aimed, directly or indirectly, at supporting this overall mission. Performance measurement in this case will confront significant problems of measurability.

- It is typical for there to be a long *time lag* between the launch of a policy and the appearance of the results intended by policy-makers. The Department of Fisheries and Oceans, for example, has launched a "policy framework" aimed at supporting the development of an economically, socially and environmentally sustainable aquaculture industry in Canada. A period of many years is likely to elapse between the launching of this policy framework and the appearance of meaningful results.

- *Attribution* of social or economic outcomes to particular public policies is almost always problematic because there is rarely a direct, one-to-one relationship between a particular policy and its intended outcome. For example, the Bank of Canada manipulates monetary policy in order to help the Canadian economy run smoothly. But many factors beyond the Bank's control – factors such as other government policies, domestic and international politics, weather, commodity prices, the economic health of Canada's major trading partners, the state of equity markets at home and abroad – have a powerful impact on the Canadian economy. The extent to which the Bank can take the credit (blame) for an improving (declining) domestic economy will therefore always be open to judgement and debate.

The challenge of measuring the performance of policy is indeed daunting, but we can make it less so by breaking down the problem into these four distinct pieces. Instead of tackling one big and ill-defined problem, we can see things in terms of four smaller and well-defined problems,

[11] At p. 14.

each of which invites a particular approach. In this part of the paper, I look at practical ways to deal with the four challenges.

B. Multiple High-level Outcomes

This is an inevitable difficulty that accompanies policy work. Policy-making is often a high-wire act. It requires a balancing of multiple forces operating in different directions. Imagine, for example, a hypothetical policy aimed at supporting a natural resource industry. It would aim, simultaneously, to support high-level outcomes related to the economic growth of the industry *and* protection of the natural environment. These two high-level outcomes are not automatically compatible! And each particular stakeholder group that is interested in the policy is likely to focus more on one outcome than on another. Owners and operators of enterprises in the industry may be most interested in the policy's economic implications. Environmentalists, as well as people who make a living from tourism, may be most interested in environmental impacts. Community leaders may be interested in the social as well as the economic outcomes of the policy. Policies therefore often aim at multiple high-level outcomes – some of which may even appear to be in conflict with one another – in order to satisfy the needs of a range of stakeholders.

This presents special problems for performance measurement because, as we have already observed, good performance measurement of a policy must rest on a foundation of consensus about the objectives of the policy. Performance measurement is meaningless in the absence of clarity about high-level outcomes.[12]

To this we can add a special complication with which many policy practitioners will be familiar. This has to do with cases where a policy's proponents will want to be *deliberately vague* about one or more of a policy's intended high-level outcomes. It may be the case – as in the example cited in "Not a Tool Kit"[13] – that the policy has important political objectives underlying the publicly stated technical objectives. Or it may be – as in the case of our hypothetical industrial policy – that the policy simultaneously requires the support of different stakeholders with different agendas. The policy's proponents may decide, based on their reading of the stakeholder environment, to publicly "play up" certain of the intended outcomes, while "playing down" certain of the others. None of this is necessarily dishonest or misleading. It may simply be the best way, under a given set of circumstances, to shepherd the policy toward the achievement of *all* of its major objectives.

Setting aside for a moment the problem of deliberate vagueness – there may be little you can do about it if you are in charge of developing a performance framework – the more fundamental problem remains. Performance measures derive their meaning from high-level outcomes. When a policy has several high-level outcomes, some of which are pulling in different or even opposite directions, how is performance measurement possible?

Your touchstone in this case will be clarity. It may be up to you as the author of the performance-measurement framework to *force* some clarity in relation to high-level outcomes.

[12] See footnote 7.
[13] See footnote 11.

The performance measurement framework that you develop for a policy, if it is to provide a basis for meaningful measurement, must be explicit in its assumptions about the high-level outcomes that the policy is supposed to be serving. As the author of the performance framework, you need not be bound by vagueness or lack of clarity that you find in the formal policy document itself. Nor should you feel yourself bound by commonly held (and, perhaps, not well articulated) perceptions of the policy that may prevail within your Department. Part of your role as the author of a performance framework is to *develop your own (well-founded) assumptions* about the intended high-level outcomes of the policy. You may be faced with a policy that lacks a clear, formal definition of high-level outcomes; or, it may be associated with several high-level objectives, with no indication of priority. Your job – as least as you develop the first draft of a performance framework – is to impose some results-based clarity onto this apparently unclear state of affairs. Based on your own knowledge, on document reviews and on discussions with colleagues, develop your own hypothesis about what would make sense as a set of high-level outcomes for the policy.[14] You really have very little choice when developing a performance measurement framework but to be explicit in your assumptions about which high-level outcomes the policy is supposed to be serving. These assumptions provide the foundation upon which to construct your performance framework. Without them, your framework is built on sand.

C. Measuring the "Unmeasurable"

It is not uncommon for public policies to aim for intangible outcomes. Policies sometimes aim, for example, to raise awareness about a particular issue, or to change widely held attitudes or perceptions. These are sometimes referred to as "soft" outcomes, a term which may suggest that they are somehow less worthy or less valid than "hard" (quantitative, objectively observable) outcomes such as increased production, higher average test scores, or reduced incidence of a disease. The reality of course is that furthering the public interest – the ultimate purpose of government – requires the pursuit of a broad range of hard *and* soft outcomes. Neither is necessarily more nor less valid than the other. Our bias toward valuing hard over soft outcomes is often related to the fact that it is simpler to measure the achievement of hard outcomes than soft ones.[15]

When faced with intangible outcomes that are important in relation to the policy, you really have no option but to try to find some way of incorporating them into your performance-measurement framework. "We can't measure *that* " is not an acceptable response if the outcome in question relates to a significant element of what the policy is trying to achieve. If important outcomes are missing from your framework, your "performance story"[16] will be seriously incomplete.

---

[14] You would do so bearing in mind that your performance framework is likely to go through several rounds of preliminary review. If others disagree with your assumptions, there will be ample opportunity to alter them! But using your assumptions as a basis for forcing a discussion about intended outcomes will be valuable, in and of itself.

[15] Our tendency to value most highly outcomes whose measurement appears to be easiest is well documented in the management literature.

[16] "Not a Tool Kit" argues that good performance measurement is an exercise in storytelling. A well developed performance framework allows you to tell a convincing story, backed by credible evidence, about the value added to Canadian society by your policy or program. See p. 1.

There are two general approaches that you can take to develop credible measures in relation to the achievement of intangible outcomes. In either case, you will try to reduce the qualitative elements of the outcome to something that can be analyzed quantitatively.

The first approach is the direct approach, where you focus on the outcome itself. Let's take the example of "partnership". The federal government is putting increasing emphasis on having government departments and agencies form policy or program delivery partnerships with entities in the not-for-profit or private sectors. Although "partnership" is not normally a high-level outcome in its own right, it is usually seen as important contributor to high-level outcomes. The reasoning supporting partnerships is normally along the lines of: *"If Department 'A' is able to form and maintain productive partnerships with key organizations in the not-for-profit sector, then the likelihood of achieving the Department's policy and program goals will be increased."*

In terms of the "logic model" discussed in "Not a Tool Kit"[17], "forming and maintaining productive partnerships" would be an important intermediate outcome that is presumed to lead to high-level outcomes. Given the importance of partnerships to the Department's business model, it would be necessary to develop a performance indicator for this intermediate outcome. How to do so?

The *direct* approach says that we focus on the outcome itself, and try to extract from it something that might be amenable to quantitative analysis. One possibility would be to undertake a regular survey of the partner organizations that work with Department 'A', systematically gathering their views on the productivity and overall quality of the partnership. In order to allow for quantitative analysis, respondents could provide their answers on numeric scales (e.g. "On a scale of '1' to '5', where '1' equals "not at all productive" and '5' equals "highly productive", how would you describe your partnership with Department 'A' ?").

The *indirect* approach would have you look at an indicator that is related to the outcome in question, but is not a direct measure of the outcome itself. This is referred to as a "proxy" indicator, because it stands in place of the thing that we wish to measure. We might choose to measure the proportion of partnerships that are renewed beyond their agreed termination date. Our reasoning here would be that if both partners agree to renew a partnership, their decision must to an important degree be motivated by a shared perception that the partnership is productive. If a high proportion of partnerships are renewed, we can (on the basis of this reasoning) conclude that the partnerships are by and large productive. It would also be important to observe changes over time in the proportion of partnership agreements that are renewed.

Inevitably, something will be lost in the translation from qualitative outcome to quantitative measure. The measures will never be a perfect reflection of reality. Indeed, if they are not carefully designed or interpreted, they may produce a distorted picture. For example, a decline in partnership renewals may be a result of cutbacks by funding agencies, and have nothing to do with the productivity of partnerships. High ratings of partnership quality that might emerge from a survey of partners could be attributed to the fact that the partners are heavily dependent on Department 'A' for funding, and therefore may be reluctant to give unfavorable ratings.

[17] Page 11.

Expect therefore that others may challenge the validity of your indicators. This is normal, and will produce a healthy dialogue that will inevitably strengthen your performance-measurement framework. Just be prepared to explain the reasoning behind your choice of indicators. Be clear in your own mind about why you think the indicator is a fair representation of reality, and why you think that the risks of producing a false or distorted picture of reality can be minimized. (See Box 1[18])

Box 1 – Measuring Professors

The problems of measuring performance in relation to intangible outcomes are not confined to public programs and policies. Academia provides an interesting case. How "good" is a professor? How can you tell if a professor is performing well?

Professors do two things – teach and produce publishable research. Teaching and research outcomes are largely intangible and qualitative. What to do?

Both direct and proxy measures can be used to address this performance measurement challenge. *Direct measures* have been used in relation to teaching performance. It is common at universities for students to complete evaluation forms where they assign numeric values to aspects of a professor's teaching capacities. The results are aggregated into an overall grade for the professor.

Proxy measures have been used in relation to the quality of a professor's research output. "Citation surveys" produce a count of the number of times that a given publication by a given professor has been cited in other publications. The assumption is that a high number of citations indicates the quality and usefulness of the research. (Others – it is assumed – wouldn't be citing the article if it wasn't any good.)

Both measures could be challenged. Student ratings might be influenced by extraneous factors. Perhaps a student found the subject matter uninteresting, even though the professor was a good teacher. Or a student's rating of a professor might be affected by a personal conflict that had nothing to do with teaching ability. Similarly, it is conceivable that a bad publication might receive many citations, because many academics might be publishing articles that criticize it!

Despite their weaknesses, student surveys and citation surveys have emerged as standard tools for measuring the performance of professors. The key to using such indicators – which aim to quantify qualitative outcomes – is to recognize both their strengths and their weaknesses. Be prepared to explain and justify your reasons for using them. When used appropriately, they will provide a useful – though never perfect – picture of reality.

[18] I thank Julian Roberts, Professor of Criminology at the University of Ottawa, for suggesting the example of citation surveys.

D. Time Lag and Attribution[19]

Between the idea
And the reality
Between the motion
And the act
Falls the Shadow[20]

i. Mind the Gap

In life as in policy, there is a *gap* between "the idea and the reality"; between what we would like to have happen (our desired high-level outcomes) and what will actually happen. For the policy practitioner, the gap is defined by the challenges of time-lag and attribution.

The gap that concerns us here has both a causal and a temporal dimension. To put it in plainly:

- you can't always get *what* you want;
- even if you get what you want, you can't always be entirely sure of *why* you got it; and
- even if you get what you want and you know why you got it, you won't always get it *when* you want it.

Causal and temporal gaps are a natural feature of government activity. Governments invest time and resources *now* in the development of policies and programs whose impact is supposed to be felt *later*. They operate on the uncertain expectation that today's investment will lead to *future* social and economic benefits. Causal and temporal gaps intercede between today's investment and tomorrow's desired outcome.

The *causal gap* creates uncertainty on two levels. First, no matter how good your planning and analysis is, you can rarely be certain that today's policy inputs and outputs will in fact lead to the desired social and economic outcomes. Most of the high-level outcomes targeted by government policies are affected by many factors outside of the government's control. Second (this is a corollary of the first), even if the desired outcomes occur, we can rarely be certain that the government's intervention (as opposed to other unrelated factors) was the primary cause.

Above, we cited the Bank of Canada's use of monetary (interest-rate) policy to support the outcome of a smoothly-running Canadian economy. The Bank's economists have theories about the links between monetary policy and Canada's economic performance. They base their policy interventions on these theories. But they also know that many powerful factors other than their own manipulation of interest rates will affect economic performance. Desired economic performance may or may not occur as a result of the Bank's policy intervention. Even if it does occur, there will never be certainty about the degree to which it was caused by the Bank.

---

[19] See "Not a Tool Kit", pp. 20 and 23-26 for a detailed discussion of time-lag and attribution.
[20] From *The Hollow Men*, by T.S. Eliot (1925).

Does this mean that performance measurement is impossible under these circumstances? No. But it does mean that we will be dealing with questions of interpretation and probability rather than certainty.

The *temporal gap* accentuates the challenges created by the causal gap. The longer the lag between the policy intervention and the appearance of the desired high-level outcomes, the greater the opportunity for the desired outcomes to be influenced by factors extraneous to the policy itself. For example, it is normally expected that changes in monetary policy will take approximately nine months to begin to have an impact on the economy. This is a relatively short time lag, but it is long enough to allow other factors to intervene in a way that may upset the Bank of Canada's calculations. The price of oil might rise or drop sharply, the United States might implement a trade-policy decision that affects markets for key Canadian exports, the stock market might crash, or a war might be declared somewhere in the world.

The temporal gap also creates an immediate performance-reporting problem. It may take many years for a policy to show evidence of having produced its intended high-level outcomes. On the other hand, performance-reporting occurs on an annual cycle in virtually all government departments and agencies. In the example noted above of the Department of Fisheries and Oceans, the Aquaculture Policy Framework operates on a long time horizon. It would not be reasonable to expect concrete evidence within a year's time that it had helped create an economically, socially and environmentally sustainable aquaculture industry And yet the Department will be required to report on performance every year. *The production of high-level outcomes and the production of performance reports occur on vastly different time-cycles!*

ii. Stand Back, and You'll See Better

To reiterate: the challenges of time-lag and attribution result from a gap between the actions we take now and outcomes that we expect to occur later. The gap occurs in time (outcomes may take many years to appear) and in causation (many factors, apart from our policy intervention, affect the outcomes). Dealing effectively with the challenges of time-lag and attribution requires that we bridge the gap, even if we can't close it. To do so, we need to see the gap in a new way. It you stand too close, the gap may appear unbridgeable. But if you stand farther back, not only will it appear to be smaller, but you will also see it in context and be able, therefore, to develop strategies for addressing it.

Standing back – getting the perspective that we need in order to solve this problem – requires as an *absolute pre-requisite* that you be clear about two things:

- what do you want to achieve over the long term (what high-level outcomes are you seeking)? and

- what are the steps by which you believe you will to get from where you are now to where you want to be?

The logic model that you must develop as a basis for a performance-measurement framework is the instrument that will force you to develop clear answers to these questions. (Readers who feel

they are not well familiar with the structure, purposes and uses of a logic model should refer to the discussion in "Not a Tool Kit."[21]) Once you have developed your answers to these questions, you can then proceed to construct a performance measurement framework that will help bridge the temporal and causal gaps.

iii. Performance Measurement vs. Evaluation

As a general proposition, you will have two instruments that will allow you to say something about the performance of your policy. One – the focus of this paper – is *performance measurement*. The other – which I will address only in passing – is *evaluation*, a distinct but related technique. Both can help you tell a story about performance and contribute to increased policy or program effectiveness. Both work from some common data sources. Both take as their fundamental point of reference the logic model that you have developed for the policy.[22] They differ, however, in their time horizons, their assumptions and their particular uses.

Performance measurement is primarily about the "here and now". It is descriptive. It looks at where things are today and asks "how are well are you doing?" in relation to the high-level outcomes you have set for yourself. It looks for evidence that you are moving in your intended direction. It helps managers make mid-course corrections to policy/program implementation, and provides a basis for being accountable to stakeholders.

Evaluation takes a longer-term perspective, looking back over a period of years at the performance of a policy or program. While performance measurement might be a basis for speculation on whether a policy was likely achieve its objectives, evaluation is more definitive. Based on in-depth research and analysis, it attempts to develop firm conclusions about whether the policy has achieved its intended outcomes. It pushes the analysis more deeply, asking whether the high-level outcomes themselves were well chosen[23]. (Performance measurement takes high-level outcomes as given.) Evaluations, unlike performance measurement, may also examine alternative ways to have pursued a given set of high-level outcomes,. As well, evaluations will normally attempt to explain *why* a policy or program has or has not achieved its objectives. They will probe the related question of whether the high-level outcomes, if achieved, were in fact caused by factors other than the policy itself. Evaluation feeds into higher-level decisions about the choice and design of policies and programs, while performance measurement is used mainly for day-to-day management and accountability.

Good performance measurement is like well-informed journalism: it can be produced relatively quickly, focuses on today, meets our immediate information needs, and helps us make rough-and-ready judgments. Good evaluation is like well researched history. It takes longer and is costlier than performance measurement, has a longer time horizon, and helps us draw deeper conclusions decisions about the nature and direction of public policy. [24]

---

[21] See pp. 11- 14.
[22] "Not a Tool Kit," p. 13.
[23] See footnote 9.
[24] See also "Performance Measurement and Evaluation. Definitions and Relationships," Washington, DC: United States General Accounting Office, April 1998. www.gao.gov/special.pubs/gg98026.pdf

iv. Practical Example 1

To apply these ideas in practice, I'll continue with the parenting metaphor. I use a silly example, but one which serves to make a point.

Let's assume that I have one child, and my desired long-term outcome for her is that she win a Nobel Prize as an adult. Let's assume that I have done research on the best way to ensure that a child grows up to win a Nobel Prize. This research tells me that making my child feel loved and building her sense of self esteem helps contribute to her intellectual development. This, in turn, increases the probability that she will win a Nobel Prize. Other research tells me that regular hugging and praising of my child will make her feel loved and will build her self-esteem.

In view of my desired high-level outcome, and given the results of my research, I institute the *Love and Self-esteem Policy*. It states that "I will pay special attention to making my child feel loved and building her self esteem." I implement the policy by hugging and praising my child as often as possible.

The logic, or the theory, behind the policy and its related implementation is:

- I hug my daughter and praise her often;
- As a result of regular hugging and praising, my daughter feels loved and has a strong sense of self-esteem;
- A strong sense of feeling loved and of self-esteem advances my daughter's intellectual development;
- Advanced intellectual development leads to my daughter eventually winning a Nobel Prize.

The four bullets constitute a simple logic model (Figure 1) for the *Love and Self-esteem Policy*. There are two especially important points to note about this logic model. (Both of these points are addressed in detail in "Not a Tool Kit".). The first is that it has been built from the top down, not the other way around. In other words, it was founded on a clear understanding of the desired ultimate outcome (a Nobel Prize). The rest of the structure was articulated by working backwards in a logical fashion to the immediate actions that could be taken today (hugging and praising) in order to achieve the ultimate outcome. This is a more significant point than it may seem at first. A common mistake in the development of logic models is to begin at the bottom and work up. (See "Not a Tool Kit" for more on this point.)

The second point is that the indicators which will provide a basis for measuring the performance of the *Love and Self-esteem Policy,* will be derived directly from the logic model. (See "Not a Tool Kit" for details on the link between the logic model and performance measures.)

My immediate problem is that I want *now* to measure the performance of the *Love and Self-esteem Policy*, but it may be 30 years or more before my daughter wins a Nobel Prize. This is the "time lag" challenge. How do I handle it?

I focus on the things that I can measure *now*: the immediate outputs, as well as the short and medium term outcomes of my policy.[25] Am I hugging/praising my child every day? Is there evidence that my child feels loved? Is there evidence of improved self-esteem and advanced intellectual development? Even though I can't possibly know today whether my child will win a Nobel Prize, this is all I have to go on right now. By measuring these things, I get an *indication* of whether my performance today is leading me in the direction of achieving my high-level outcome. (Hence the term "performance indicator".) I can't know for certain if I will get to where I want to go, nor can I be certain that even if I *do* get to my ultimate outcome it will have been as a result of my policy. But based on the information available to me now, I can make informed judgements about whether or not my performance today appears to be leading me in the right direction.

I use my theory – my logic model – to connect what I am doing now with my intended long term outcome. The logic model gives me a basis for telling a believable story about my performance, even though it is unknown whether the ultimate outcome will occur. My "performance story" would be structured in the following way:

> *I have a sound basis for believing that regular hugging and praising of my child will make her feel loved and will build her self-esteem. I also have a sound basis for believing that if my child feels loved and has high self esteem, her chances of winning a Nobel Prize as an adult will be vastly improved. Given what I can demonstrate to you now regarding my record of hugging and praising my child, and given what I can demonstrate now regarding the impact of this activity on her sense of feeling loved and on her sense of self esteem, I submit that I am performing well because I am contributing to the eventual achievement of my desired ultimate outcome.*

The challenges of time-lag and attribution mean that I cannot directly measure my performance against the achievement (or not) of the ultimate outcome. But I can measure performance against evidence that I am moving in the right direction. It then becomes a matter for the parties who receive my performance reports to make their own judgements about the validity of my reporting. Other parties will indeed be able to make informed judgements because:

- I have *revealed my performance-measurement assumptions* by developing and explaining the logic model that underlies the policy; and

- I have been clear about the long-term objective that is the foundation of my logic model.

Now, let's move ahead in time, 40 or 50 years or so. By then, either my child will have won a Nobel Prize or it will have become obvious that she will not. Armed with this certain knowledge about the attainment (or not) of my ultimate objectives, what can I say about my performance?

[25] See p. 10 of "Not a Tool Kit" for a discussion of performance-measurement terminology.

If my daughter won the Prize, does that mean that I have performed well? If she didn't, does that mean that I have performed poorly?

The answer is not obvious. I might have implemented the *Love and Self-esteem Policy* to perfection, and yet my daughter might not have won the Nobel Prize in any case. Conversely, I might have done a poor job with the policy, and yet my daughter won. Performance measurement, being mainly a descriptive exercise, will not get us far in solving the attribution problem, i.e. drawing the link between "good performance" and achievement of the ultimate objective. It may give us a reasonably good indication of whether or not our intervention has been effective in terms of producing outcomes, but it cannot do more than that. (Having said that, we should also note that most of the time, a "reasonably good indication" will be all that we need.)

If we want a more definitive reading on attribution, it would be necessary to conduct an evaluation. As suggested above, in an evaluation we would take a detailed (and more costly and more time-consuming) look at questions such as:

- was the research that underpinned my assumptions valid?
- how effectively did I implement the policy (did I hug well? praise well?)
- does it appear that factors apart from my policy had a significant impact on the ultimate outcome? would the same outcome have occurred without my intervention?

v. Practical Example 2

Now, let's apply the same kind of reasoning to a real federal government policy. The Department of Fisheries and Oceans (DFO) recently published[26] a detailed description of a new policy, the "Aquaculture Policy Framework" (APF). The document provides a basis for us, in this paper, to consider how we might apply the principles of RBM and performance measurement to a government policy.

In examining the published material, it is possible to discern two *high-level outcomes* to which DFO wants to contribute through the APF:

- the development of a "durable" Canadian aquaculture industry, i.e. one that is sustainable from a commercial, environment and social perspective;
- the development of an Canadian aquaculture industry that is regarded internationally as a world leader.

[26] "DFO's Aquaculture Policy Framework," Ottawa: Fisheries and Oceans Canada (Communications Branch), 2002. DFO defines aquaculture as "the farming of aquatic organisms in marine or fresh water." I became aware of the APF in the course of providing advisory services to DFO. However, all of the material related to the APF that is presented in this paper is either contained in or could reasonably be inferred from public documents.

In order to contribute to these outcomes, the APF provides a framework within which a wide range of *outputs* are supported and encouraged. These include (but are not limited to):

- aquaculture-related research;
- industry development programs;
- enhancement of DFO's internal capacity related to aquaculture (e.g. through resources allocated to aquaculture);
- communication with stakeholders to ensure broad ownership and understanding of the goals of the APF, and deeper understanding (particularly among aquaculture operators) of DFO's regulatory responsibilities;
- appropriate regulatory and administrative frameworks related to aquaculture;
- support for equitable, predictable and timely access to aquaculture sites
- internal communication and awareness building related to the importance of aquaculture in relation to DFO's overall mission.

These outputs could be expected to contribute to the achievement of *first-level (immediate) outcomes* that might include:

- shared understanding between DFO and external stakeholders (e.g. the aquaculture industry) regarding the objectives and implications of the APF;
- an enabling environment (including the regulatory environment) that is supportive of the development of the aquaculture industry;
- compliance by aquaculture industry operators with the regulations and administrative requirements overseen by DFO;
- responsible management practices by aquaculture operators; i.e. they pay heightened attention to issues such as food safety, navigational safety, environmental sustainability and social cohesion;
- cultural change in DFO, leading to widespread acceptance of aquaculture as an important and legitimate area for the Department, alongside its traditional focus on the wild fishery.

These first-order outcomes could be expected to contribute to the achievement of *second-level (intermediate) outcomes* that might include:

- an aquaculture industry that is internationally competitive, by virtue of access to skills, knowledge, technology, capital and markets;

- public confidence that aquaculture products are safe to eat, that aquaculture is not harmful to the environment and that it is not crowding out other users (e.g. recreational) of aquatic resources.

- social cohesion, in the form of a positive impact by aquaculture on communities (with a particular emphasis on First Nation communities);

- environmental sustainability – aquaculture is conducted in a way that minimizes harmful impacts on the natural environment.

These four second-order outcomes could be expected to contribute to the high-level outcomes identified above.

What we have described here is a *logic model* for the APF (see Figure __) that leads us from outputs through first-level and second-level outcomes to high level outcomes. Exactly as was the case with the hypothetical *Love and Self-esteem Policy* (above), our logic model provides us with a framework within which to approach the problem of performance measurement.

Developing Performance Measures. Most importantly, the logic model provides a basis for developing a set of performance measures[27]. One you have done the hard work of developing a logic model that shows how the policy is expected to contribute to ultimate outcomes, the task of identifying performance measures becomes relatively simple. Table 1 provides examples of the link between the APF logic model and a set of performance measures.

Table 1

Layer of Logic Model	Item in Logic Model	Example of a Related Performance Indicator
1st-level Outcome	***Enabling environment***	**Awareness of rules, regulations, etc. related to aquaculture**
1st-level Outcome	***Compliance***	**No. of prosecutions in relation to total aquaculture operations**
2nd-level Outcome	***Competitiveness***	**World-wide market share of Canadian aquaculture**
2nd-level Outcome	***Environmental Sustainability***	**Habitat quality in aquaculture zones**
2nd-level Outcome	***Public Confidence***	**Public attitudes re aquaculture (from surveys)**
Ultimate Outcome	***"Durable" Industry***	**Landed value of fishery products, wild vs. cultured**
Ultimate Outcome	***"Durable" Industry***	**Investment in aquaculture**

We want to be able to use our performance measures, together with our logic model, to help us address three kinds of questions that we and our stakeholders care about very much:

- does the APF appear to be generating the kinds of immediate and intermediate outcomes that we anticipate in the logic model?

[27] See "Not a Tool Kit", pp. 12, 15, 16.

- as a result, are we increasing the likelihood that the APF is contributing to the desired ultimate outcomes?

- can we tell a believable story, backed by credible evidence, that implementation of the APF is moving in the right direction?

The Four Challenges. Notice that all of the four challenges noted above – multiple high-level objectives, measurability, time-lag and attribution – apply when we attempt to measure the performance of the APF. By breaking the APF into its component parts – by exposing its "anatomy" – the logic model helps us identify where these challenges reside. This relatively straightforward process of identifying the critical performance measurement challenges, which is facilitated by the logic model, helps us get a grip on how to begin addressing the challenges.

Multiple high-level objectives. We find these, for example, in the first-level and second-level outcomes. Within the first-level outcomes, we see both "enabling environment" and "responsible management". There might be circumstances in which these two outcomes might be pulling in opposite directions. One might interpret "enabling environment" as implying a regulatory regime that allows the maximum possible freedom for operators. This might be inconsistent with an emphasis on "responsible management." Similarly, with respect to the second-level outcomes, there could be conflict between the emphasis on competitiveness, and, on the other hand, the emphasis on environmental and social sustainability. While the logic model itself does not provide answers to the question of how to balance these tensions, it does make it easy to recognize that the tensions exist. It does help lead us to the conclusion that it is critical to measure and report on social as well as economic performance, so that we get a well-balanced picture of the APF's performance.

Measurability. An outcome such as "public confidence" creates the kind of measurement challenge discussed above at pp. 12-14. It is a "soft" outcome. We want to try to reduce it, either by a direct or a proxy (indirect) approach, to quantitative terms. The indicator shown in Table 1 – public attitudes regarding aquaculture derived from surveys – is an example of the direct approach. People would be asked direct questions about their confidence in the aquaculture industry, and the resulting data would be quantified. We could also take a proxy approach by way of another possible indicator for "public confidence", which is "value of sales in Canada of aquaculture products." This is an indirect, quantitative measure of public confidence. The reasoning behind this indicator is that sales in Canada of aquaculture products are at least party attributable to the Canadian public's confidence in the industry. If, for example, Canadian consumers were to have serious concerns about the safety of eating aquaculture products, this would presumably be reflected in sales figures.

Time Lag and Attribution. Measurement of the APF's performance in contributing to the ultimate outcome – creation of a "durable" Canadian aquaculture industry – provides a good example of the challenges of time-lag and attribution. Many factors apart from the APF will have an effect on the durability of the aquaculture industry in Canada; as well, it may be some years before we can safely conclude that aquaculture has established itself as a durable industry. We can deal with these challenges by emphasizing performance indicators at the level of immediate (first-level) and intermediate (second-level) outcomes. We focus on what it makes

sense to measure *now*. These indicators measure outcomes over which the APF has some reasonable (albeit less than 100 percent) level of control and which are capable of showing meaningful change over the course of the typical one-year reporting cycle.[28] Even though it may not be possible to measure directly the performance of the APF against the ultimate outcomes, it will be possible to measure the APF's performance against certain immediate and intermediate outcomes, and then make an argument that the APF's performance at the immediate or intermediate-outcome level provides credible evidence that the APF is contributing to the intended ultimate outcomes.

One example of a performance measure at the intermediate outcome level that would serve this purpose is "habitat quality in aquaculture zones". The APF – through its support for regulation and research and awareness raising – has some reasonable degree of control over this outcome. It is also an outcome that can be expected to show meaningful change on an annual basis. Measurement of performance against this outcome can be used to develop the following kind of argument:

> *If we can demonstrate that we have been successful in contributing to safeguarding and/or improving habitats in aquaculture zones, then, given the assumed linkages in the APF logic model, it is reasonable to conclude that we are contributing to the ultimate outcomes, i.e. building a durable industry and building a reputation for the industry as a world leader.*

vi. The Link with Evaluation

As was observed above (pp. 17-18), it is important to understand that performance measurement has its limitations. The challenges of time-lag and attribution will make it impossible to measure the APF's performance directly against the achievement of ultimate outcomes. The old saying that "the perfect is the enemy of the good" applies here. Sometimes a less-than-perfect instrument is, under the circumstances, the best one for the job at hand. Performance measurement is indeed a "second-best" instrument – but a very useful instrument nonetheless. If you make judicious use of first and second-level outcome measures in the context of a sound logic model, then performance measurement can help you assess the likelihood that that the APF is contributing to its intended ultimate outcomes. But if you want a more definitive understanding, it would be necessary to undertake an evaluation – initially after several years had passed, and then at similar intervals over the life of the policy. It is only through the in-depth analysis that is characteristic of an evaluation that you would be able with a relatively high level of certainty and credibility to make claims about the impact of the APF on ultimate outcomes.

vii. Linking the Present with the Future

Whether you want to measure the performance of a policy or a program, the goal is the same: find a credible and convincing way of linking the present with the future.

---

[28] See "Not a Tool Kit" at pp. 25-27 for a discussion of two types of indicators – those that are a "fair reflection" of program/policy performance (suitable to be used as performance indicators) and those that are "related to" policy/program performance (not suitable to be performance indicators, but still an important part of the performance reporting package).

Whether you are involved in the implementation of a program or a policy, the same basic set of circumstances applies: *you are committing time and resources today for outcomes in the future that are uncertain.* In the example of the APF, we have a policy framework in place today that is intended to yield beneficial outcomes in the future. The outcomes are well defined now, but from where we sit today it is far from certain that they will materialize. And even if they should materialize, it will often be difficult to discern immediately whether they have been caused by the APF or by other factors.

Citizens have no less a right to be informed about the performance of policies than of programs. In order to explain and justify the allocation of resources to the APF or any policy (or program) you need to have a way of connecting what you are doing now with where you want to be in the long term. This connection needs to be clear and must make sense not only in the minds of the people responsible for the policy, but also in the minds of external stakeholders (citizens, civic groups, private sector operators, politicians, etc.).

Performance measurement helps you make that connection. It helps you tell a believable and compelling story about *why* a policy was conceived in the first place, and *whether or not* it appears to be on the right track.

This paper has tried to make the point that it is no less feasible to do this for policies than for programs. In either case, the thought process is identical. The argument that "performance measurement does not apply to policies" simply does not hold water!

Figure 1

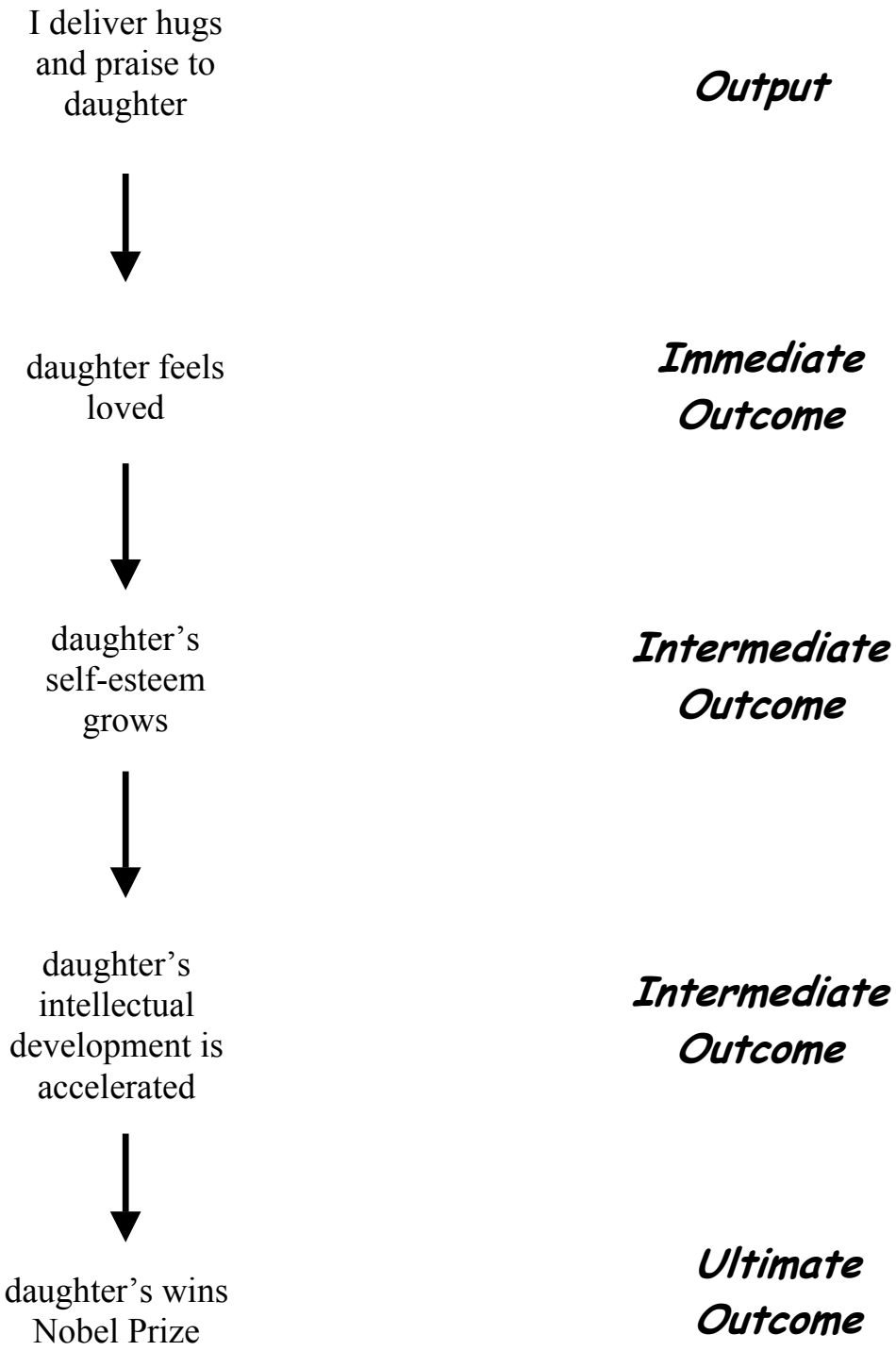
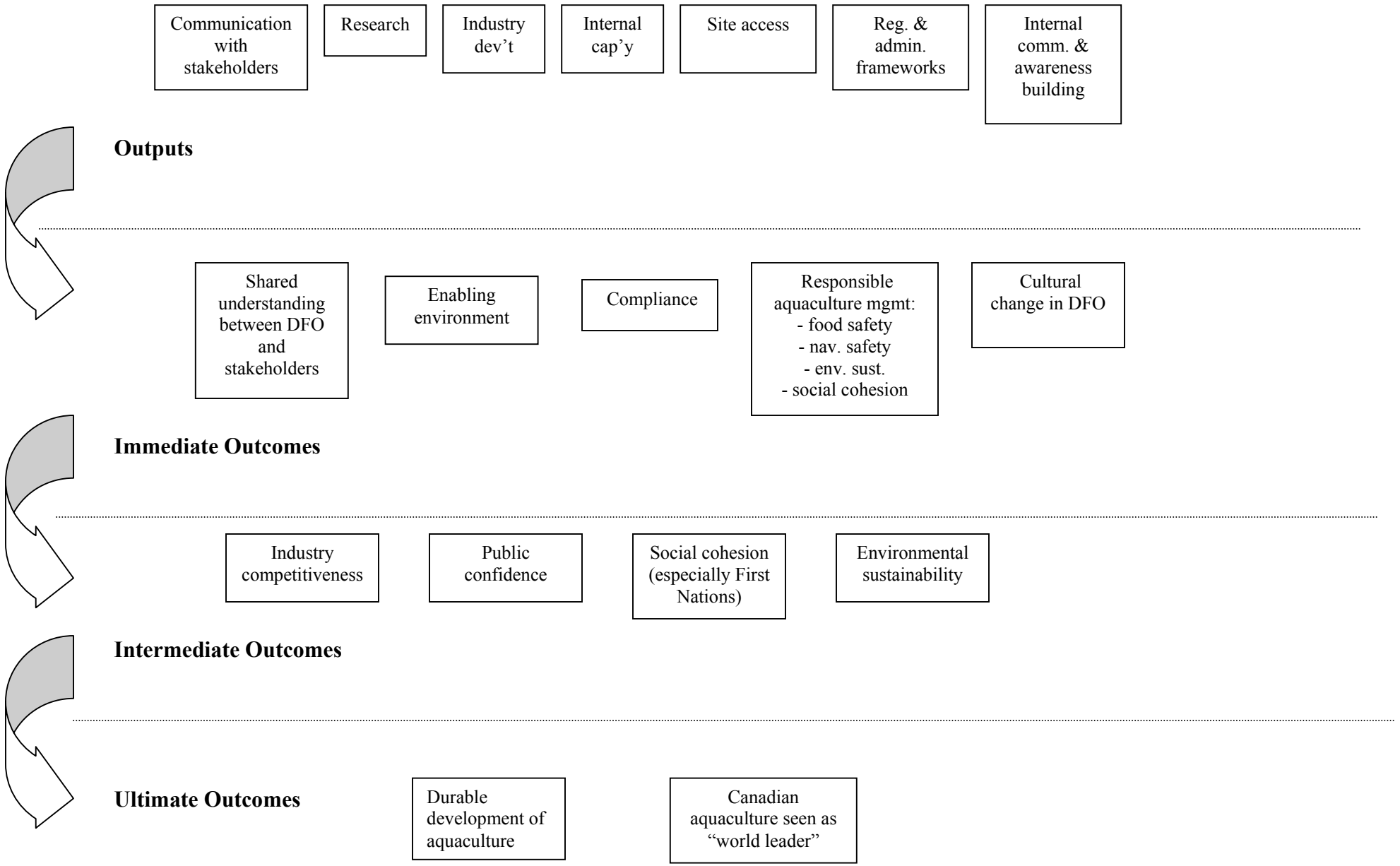

I deliver hugs and praise to daughter
Output
daughter feels loved
Immediate Outcome
daughter's self-esteem grows
Intermediate Outcome
daughter's intellectual development is accelerated
Intermediate Outcome
daughter's wins Nobel Prize
Ultimate Outcome

Figure 2


Communication with stakeholders
Research
Industry dev't
Internal cap'y
Site access
Reg. & admin. frameworks
Internal comm. & awareness building
Outputs
Shared understanding between DFO and stakeholders
Enabling environment
Compliance
Responsible aquaculture mgmt:
- food safety
- nav. safety
- env. sust.
- social cohesion
Cultural change in DFO
Immediate Outcomes
Industry competitiveness
Public confidence
Social cohesion (especially First Nations)
Environmental sustainability
Intermediate Outcomes
Ultimate Outcomes
Durable development of aquaculture
Canadian aquaculture seen as "world leader"

The Institute On Governance (IOG) is a non-profit organization founded in 1990 to promote effective governance. From our perspective, governance comprises the traditions, institutions and processes that determine how power is exercised, how citizens are given a voice, and how decisions are made on issues of public concern.

Our current activities fall within these broad themes: building policy capacity; Aboriginal governance; accountability and performance measurement; youth and governance; citizen participation; governance and the voluntary sector; and technology and governance.

In pursuing these themes, we work in Canada and internationally. We provide advice on governance matters to organizations in the public, private and non-profit sectors. We bring people together in a variety of settings, events and professional development activities to promote learning and dialogue on governance issues. We undertake policy-relevant research, and publish results in the form of policy briefs and research papers.

You will find additional information on our themes and current activities on our web site, at www.iog.ca.