

REPORTING ON OUTCOMES: SETTING PERFORMANCE EXPECTATIONS AND TELLING PERFORMANCE STORIES

John Mayne
Office of the Auditor General of Canada
Ottawa, Ontario

Abstract: Results, and more particularly outcomes, are at the centre of public management reform in many jurisdictions, including Canada. Managing for outcomes, including setting realistic outcome expectations for programs, and credibly reporting on what was achieved are proving to be challenges, perhaps not unexpectedly, given the challenges faced in evaluating the outcomes of public programs. This article discusses how the use of results chains can assist in setting outcome expectations and in credibly reporting on the outcomes achieved. It introduces the concepts of an evolving results-expectations chart and of telling a performance story built around the program's results chain and expectations chart.

Résumé: Les résultats, et plus particulièrement les conséquences, sont au cœur de la réforme de la gestion publique dans de nombreuses administrations incluant au Canada. La gestion axée sur les résultats, notamment l'établissement d'attentes réalistes en matière de résultats pour les programmes, et la communication d'une information crédible sur ceux-ci s'avèrent des défis, ce qui n'est peut-être pas étonnant compte tenu des défis que pose l'évaluation des conséquences des programmes d'intérêt public. Le présent article analyse la façon dont l'enchaînement des résultats peut servir à établir les attentes et rendre compte, de façon crédible, des conséquences atteintes. On y présente l'idée d'un tableau exposant des attentes évolutives face aux résultats ainsi que celle d'une description du rendement centrée sur l'enchaînement des résultats du programme et le tableau des attentes.

“Results” and “performance” are becoming mainstays of discussions of public management. Managing for results and reporting of results achieved are part of the lexicon of public sector management. Public managers are being asked to focus on producing

Corresponding author: John Mayne, 654 Sherbourne Rd, Ottawa, ON K2A 3H3;
<john.mayne@rogers.com>

results and producing them in an economical way with the public money entrusted to them.

The basic elements of managing for results, or results-based management, can be set out quite easily. In the context of programs, managing for results requires (Auditor General of Canada, 1997, pp. 11–12):

1. fostering an organizational climate that encourages managing for results,
2. agreeing on expected results,
3. measuring results to improve performance, and
4. effectively reporting performance.

Managing for results, however, has proven to be quite difficult to implement (Auditor General of Canada, 2000b; General Accounting Office, 2003; Perrin, 2002; Thomas, in press). Measurement of results — outputs and outcomes — in the public sector is often thought of as the biggest challenge. Moving to a results-focused culture is also seen as a huge challenge, and good performance reporting seems to be painfully slow (Auditor General of Canada, 2000a). Agreeing on expected results is often considered a rather straightforward challenge. Experience with managing for results, however, shows that setting performance expectations may be the most difficult aspect of managing for results to accomplish. Wholey (1997) has pointed out, “The most important initial step in performance-based management is getting a reasonable degree of consensus on key results to be achieved” (p. 100).

The concept of “performance” requires a comparison of what was expected with what was achieved. It is not possible to assess performance either for managers or for the public without knowing first what level of performance was expected.¹

The purpose of this article is to lay out a practical approach to settings expectations for managing public sector programs, in particular outcome expectations, to discuss the idea of telling a performance story, and to suggest what members of Parliament and the public should expect in government performance reports on program outcomes.

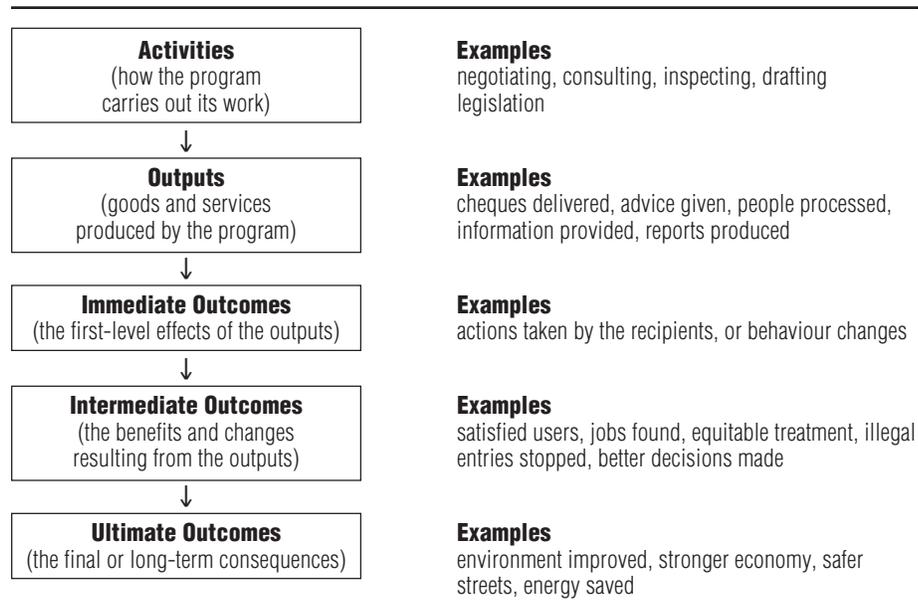
Defining Terms

Before beginning, we need to define the terms we are using. Government programs undertake a number of *activities* that produce a va-

riety of *results*. Programs deliver two kinds of results: *outputs*, the direct products and services produced by government activities, such as an unemployment cheque or some requested information; and *outcomes*, the consequences (both intended and not) of those outputs on Canadians and our society. Outputs are results that managers can largely control, while the outcomes that managers are trying to accomplish are influenced by factors outside their programs.

The final or ultimate results sought, such as general improvement in the well-being of Canadians, the economy, or the environment, are *end outcomes* (sometimes called long-term, final, or ultimate outcomes). Between the outputs and the end outcomes, there is a sequence of *immediate* and *intermediate outcomes* that are expected to lead to a desired result but are not ends in themselves, such as changes in the actions of program clients. Immediate outcomes are more easily linked to the activities of a program than are end outcomes. A *results chain* (Figure 1) shows this logical sequence of outputs and outcomes that occurs as a result of a program's activities.

Figure 1
A Results Chain



Articulating with some clarity what results a program is intended to accomplish is critical to good results management and reporting. A wide range of terms are used to describe these normative statements:

objectives, goals, strategic outcomes, expected results, planned results, targets, and expectations, to name a few. Among these, it is essential to distinguish between general statements that set the direction of the overall intent of the program and more concrete statements specifying what is to be accomplished over a time period.

Both types of statements of intentions are needed. The first type (*objectives*) sets out, at a high level, the general direction and end state sought, but it often does not specify the extent of results sought or the time frame to accomplish them. Objectives are usually set out in legislation or by governments as statements of general policy intent. Objectives link to the mission, vision, and policy goals of an organization and set the stage for the second set of statements, the more concrete *performance expectations*. Performance expectations define the specific results expected, the extent (how much is expected), and the timeframe. Specific expectations are sometimes set out in legislation but often not. Rather, it is left to the public sector managers to implement, with the approval of ministers, the intent of the objectives set by governments and/or legislatures. Good performance expectations allow one to know and determine if what has been set out to be accomplished has been achieved. Having clear statements that specify what is expected is essential to performance information. Without them, all one has is results information.

Historical Background

In the past, setting expectations for performance has normally involved setting expectations for outputs. Outputs are well understood because they represent the direct result of program activities. They are usually quite visible and measurable, and one can comfortably talk about being accountable for them because they are what programs managers and staff can, for the most part, control. For many years, the performance reporting regimes of a number of jurisdictions such as the U.K. and New Zealand were heavily based on setting specific output targets and reporting on the extent to which outputs had been produced.

Today, most jurisdictions, including the U.K. and New Zealand as well as Canada, are based on or include a strong focus on outcomes. A management and reporting regime focusing on outcomes is quite different from one focused on outputs. The production process for outcomes, be they immediate, intermediate, or end outcomes, is usually not as well understood as for outputs. The linkages between

various levels of outcomes may not be well known, and the measurement of the outcomes themselves may be quite a challenge. By definition, outcomes are not controlled, but rather programs seek to influence their occurrence by carrying out certain activities and delivering certain outputs. As a result, there is much less comfort with being accountable for outcomes because those accountable for them do not control them. Yet, despite these different conditions, the usual approach to setting expectations for outcomes has simply adopted that used for outputs by attempting to set numerical targets for each outcome specified.

This approach has proven to be less than satisfactory because it does not take into account that outcomes can be difficult to measure, links between outputs and various levels of outcomes can be difficult to establish, and being accountable for single specific numerical outcomes is often not very realistic nor, more importantly, useful. For example, it may be easy to set targets for literacy and numeracy, but improvements in those areas may be at the expense of skills that are more difficult to measure, such as creativity (*Economist*, 2001). In this article, I suggest an alternative approach that recognizes that outcomes are not like outputs.

THE RESULTS-EXPECTATION CHART

When considering the performance of a program, a logic model or results chain is often developed. The evaluation literature contains much discussion of logic models, their use and development (Millar, Simeone, & Carnevale, 2001; Montague, 2000; Rossi, Lipsey, & Freeman, 2004, pp. 95–96). A logic model or results chain is usually in the form of a diagram (see Figure 1) of how the program is supposed to work. It describes the theory behind the program: how activities undertaken are intended to produce a variety of outputs, which in turn are expected to result in a chain of subsequent outcomes.

Good performance reporting tells how well a program has done at what cost. Good management requires that a manager know when expectations have been accomplished and be able to demonstrate this. Clear and concrete expectations about the outputs and outcomes to be accomplished enable a manager to do this, and to know how successful a program has been and where improvements are required. Tracking the key activities undertaken, the outputs produced, and the resulting outcomes builds an understanding of the extent to which a results chain reflects reality and, when the planned

results are not forthcoming, where changes in activities and outputs need to be made.

A results chain should be at the heart of setting outcome expectations for a program. A sensible way of setting performance expectations is to develop a results chain. Expectations for a program are not that just one or two specific outputs and outcomes will occur, but rather that the whole chain of events presented in the results chain — the program theory — will occur. Setting expectations for managing and reporting performance entails setting out a *results-expectations chart*.

This remains the case when the high-level objectives have been set by governments or legislatures. In order to effectively implement government intentions, understanding which specific activities, outputs, and chain of outcomes are likely to best lead to the objectives is essential to good public management.

A results-expectations chart requires:

- setting out a results chain that clearly defines each key element of the chain;
- presenting the contribution or intervention logic, including the context, of why and to what extent it is expected that the activities and outputs of the program will contribute to the sequence of expected outcomes;
- identifying the clear and concrete outputs to be produced; and
- identifying at least some of the outcomes to be achieved in clear and concrete terms — it may not be useful to try to measure everything.

A results chain is like a logic model, which can be developed to show, usually in some detail, the causal sequence of outputs and outcomes. But the results chains suggested here do not try to provide that level of detail. They are based on the espoused theory of the program. As a result, they do not necessarily require confirmation about causality, which is often associated with developing a program theory on which a logic model is based. Results chains are suggested here as a structure for describing the expectations of a program (and as a basis for reporting the subsequent performance story). Of course, the better understood the theory of the program is, the more realistic the expectations set out will be.

Where the results chain describes a program, it should include the planned expenditures. Where efficiency is an important aspect of performance, the expected unit costs of the outputs produced should be part of the results chain description of outputs. A complete picture of performance expectations, in addition, would discuss the legitimacy of the expectations — how they were set, their consistency with legislation, mandate, and mission, their reasonableness, and their significance.

Table 1 outlines a generic results-expectations chart, while Table 2 provides an example of a results-expectations chart for a unit in a national audit office working to improve the practice of accountability for results. The example is typical of most government programs in that the outcomes sought occur well outside the program in question, and are at best influenced by the activities of the program.

Table 2 sets out both the underlying results chain (the espoused theory of the program) as well as the results-expectations chart. Depending on the specific case, the inclusion of both may or may not be a useful approach. The results chain shows the logic of the program, while the results-expectations chart sets out more specifically what is expected to be accomplished by the program, using the results chain as a reporting structure. Note also that the results-expectations chart is a chart of intentions, not facts or reality. Measurement is required to determine the extent to which the expectations listed in the chart do in fact occur.

Setting out performance expectations in this form requires more information than in the more traditional approach of identifying a few targets, largely because outcomes are more challenging to deal with than outputs. Public programs are usually complex interventions trying to alleviate a problem or maintain a condition. Certain aspects of the program might work and be successful (certain outcomes might have been achieved) while other aspects might not (other outcomes might not have been achieved). For still other aspects, it might not be clear what has been achieved (perhaps the links between aspects of the program and the outcomes sought cannot be demonstrated). Of course, in many programs, certain key outcomes may be identified and success readily determined in those terms. But even then, success needs to be further discussed by setting out the likely extent of the contribution made by the program to the expected outcome.

Table 1
A Generic Results-Expectations Chart

| | |
|--------------------------------|--|
| Rationale | A short statement of the problem or situation being addressed and the overall rationale for the program and how it is to address the concerns. The discussion would provide the general objectives for the program, i.e., what we expect the impact of the program will be. |
| Activities/ Outputs | <p>What inputs are being used? Why are the activities undertaken? Why are the specific outputs produced? How does this link to the mandate of the organization? What is expected to be produced this year, and how does this compare with previous years and/or other entities/benchmarks/best practices?</p> <p>Performance measures (examples)</p> <ul style="list-style-type: none"> • number of units produced • average cost of each unit <p>Planned spending for fiscal year</p> <ul style="list-style-type: none"> • \$ • FTEs <p>Fiscal year expected products (examples)</p> <ol style="list-style-type: none"> 1. number and type of products expected in fiscal year 2. average cost of product in fiscal year |
| ↓ | |
| Target Group | Who are you trying to influence? |
| ↓ | |
| Partners | With whom do you have formal or informal partnering arrangements? What role do they play in your results chain relative to your own role? |
| Immediate Outcomes | <p>Describe what you expect to occur as a direct result of these activities and products. What will be the direct effect on target groups? Describe this first-level theory of the program, explaining why you expect to have a direct impact. How will you know if you have made a difference?</p> <p>Performance measures (examples)</p> <ul style="list-style-type: none"> • measure of extent to which expected immediate outcome occurs <p>Expected immediate outcomes (examples)</p> <ol style="list-style-type: none"> 1. target for results of activities/products to occur by end of fiscal year or later 2. expected immediate outcomes of activities/products on clients |
| ↓ | |
| Intermediate Outcomes | <p>What do you expect to happen as a result of your immediate outcomes? Describe how immediate outcomes described above are expected to affect stakeholders and others who might benefit by the business line, but perhaps are not direct clients. What are the broader implications and impacts of carrying out the activities, beyond the direct benefits to clients?</p> <p>External factors What other factors affect these intermediate outcomes? How would they do so?</p> <p>Performance measures (examples)</p> <ul style="list-style-type: none"> • measures of impact of activities on clients, stakeholders, and others <p>What do you expect to achieve by the end of this fiscal year or later? Performance expectations might indicate how far you might expect to have progressed toward a particular outcome, and/or they might include future expectations.</p> <p>Expected intermediate outcome (examples)</p> <ol style="list-style-type: none"> 1. Target for impact of activities/products to occur by end of fiscal year or later 2. Expected outcomes of activities/products on indirect clients, stakeholders, and others affected by activities |
| ↓ | |

Table 1 (continued)**End Outcomes**

| |
|---|
| Describe what you ultimately expect to achieve (link to earlier objectives), and what you expect the contribution of the activities of this program to be. Also describe what other entities/groups or factors will contribute to this goal, and the nature of their contribution. How will you know when your goal is being achieved? How will you know if you have made a difference? What are the societal indicators that you would expect to influence through this program? |
|---|

Note. Not all the information needs to be squeezed into boxes. If more space is needed, additional explanations for any of the boxes can be provided in the text or footnotes.

Developing a results-expectations chart would have many of the advantages for program managers often identified with developing a logic chart or results chain. It provides a framework by which program management (and stakeholders) can engage in discussion about the rationale, objectives, and structure of the program. Developing a results chain often provides a useful tool in the development of program delivery design, monitoring, and evaluation. And it can be very useful in identifying the measures and indicators that might be used to know how well the results chain/expectations chart is being realized.

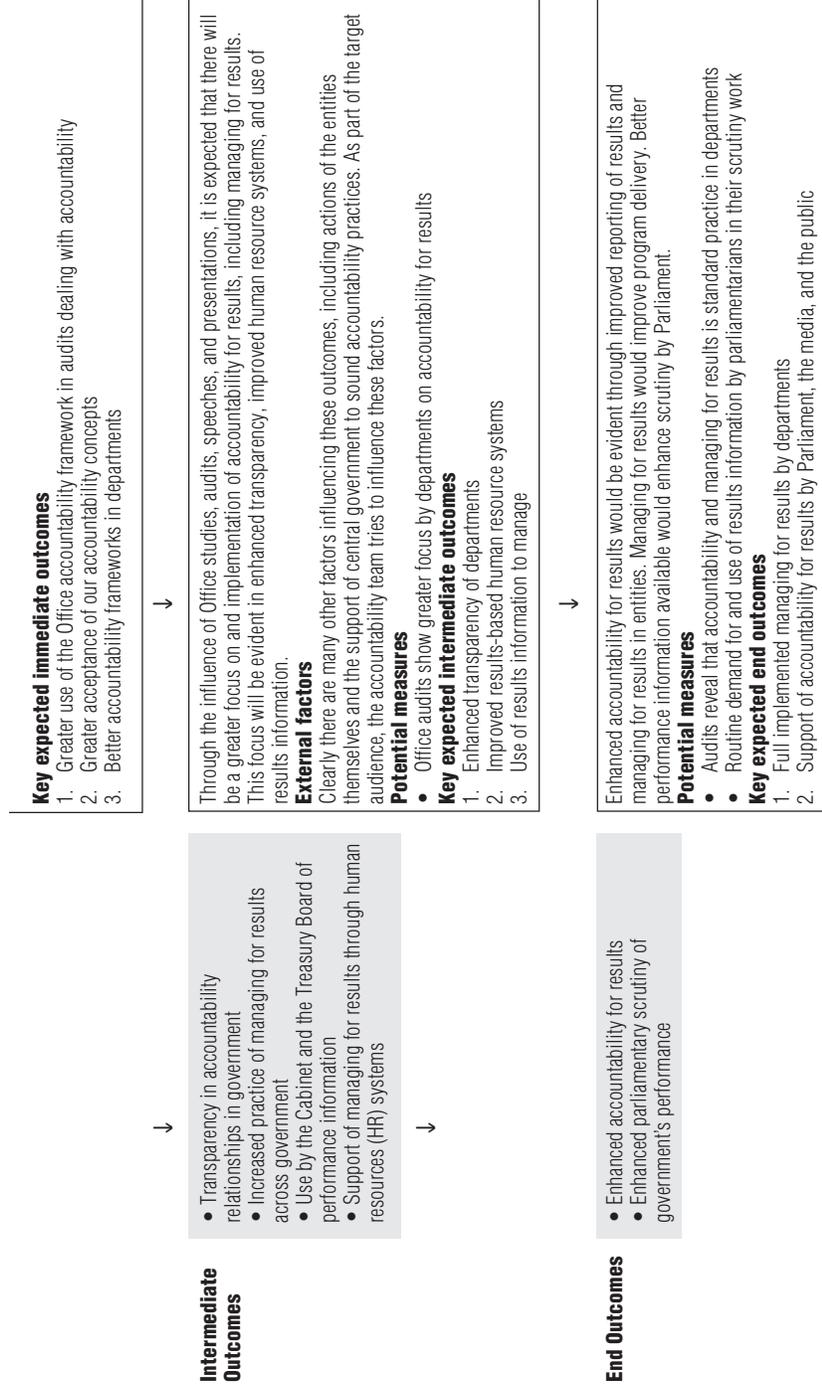
Limitations also need to be recognized. The results-expectation chart may become outdated if it is not updated as new information is acquired. And it needs to be kept in mind that a results-expectations chart is itself only a theory until evidence is acquired supporting the theory or suggesting changes. It is a means to an end.

Setting out a results chain and a results-expectations chart requires some hard work but is essential for managing for results. Managing for results requires a focus on the whole results chain or expectations chart. Too often, managing for results focuses only on the expected end outcome. But in most programs, there are many assumptions that must be realized and events that must occur if that final outcome is to be achieved (or contributed to). Without an understanding of the intended chain of events a program is trying to influence, focusing on the end outcome is not likely to be very useful.

An additional benefit of using a results-expectations chart is that, rather than relying on the more traditional one or two indicators, the concern about being accountable for outcomes should be reduced. The Office of the Auditor General of Canada (2002) has discussed

Table 2
Example of a Results-Expectations Chart

| Audit Office Activities to Enhance Accountability Practices and Managing for Results in Government | |
|--|--|
| <i>Results Chain</i> | |
| <i>The 2001–02 Expectations</i> | |
| Rationale | The Office has a long-standing interest in improving accountability practices in government, especially accountability to Parliament. Through pertinent audit reports and related communication on accountability and governance issues, the Office expects to draw attention to significant accountability issues and encourage better practices. |
| Activities/Outputs | <ul style="list-style-type: none"> • Developing frameworks (principles, audit criteria) for good accountability • Identifying weaknesses and strengths in entities • Clearly articulating accountability for results communicated in audit reports, presentations, speeches, and advice <p style="text-align: center;">↓</p> |
| Target Group | <ul style="list-style-type: none"> • Office auditors, federal departments, central agencies, and parliamentarians <p style="text-align: center;">↓</p> |
| Immediate Outcomes | <ul style="list-style-type: none"> • Audits of departments and agencies and studies examining accountability and managing for results, with consistent criteria and recommendations • Better understanding and acceptance in government of “accountability for results” • Improved accountability frameworks for departments, programs, and horizontal issues <p style="text-align: center;">↓</p> |
| | <p>With the aim of improving accountability practices in government, the accountability team develops and communicates appropriate principles and frameworks for sound accountability practices focusing on accountability for results. In addition, accountability practices in entities are identified and assessed by the team and by entity teams through studies and audits. In addition to ongoing advice in the Office and outside communication, each year a number of activities are identified to further these aims.</p> <p>Key expected products (2001–02)</p> <ol style="list-style-type: none"> 1. Audit published on accountability in new governance arrangements 2. Office guide on auditing accountability <p style="text-align: center;">↓</p> <p>There is a wide range of target groups who can influence accountability practices.</p> <p style="text-align: center;">↓</p> <p>The frameworks and principles developed are used to further a common understanding and audit approach to accountability in the Office through advice to other teams, as well as in government more generally. It is expected that the more consistent the recommendations on accountability throughout the Office, the greater will be the influence on government accountability practices. Our outreach activities are expected to lead to better understanding and acceptance in government and more broadly of the concepts of modern accountability for results.</p> <p>Potential measures</p> <ul style="list-style-type: none"> • Extent of use of the accountability framework in office audits |



the need for a new perspective on accountability, namely that accountability for outcomes should mean being able to demonstrate:

- the extent to which the objectives and expected outcomes are being achieved;
- in a reasonable fashion, the extent to which the program has influenced or contributed to those accomplishments, namely to what extent the program has made a difference;
- what has been learned; and
- that the means used were proper.

This perspective recognizes that outcomes are influenced by other factors and looks to managers to maximize the possibility that the services and products they do control adequately contribute to the end outcome sought. Setting out expectations in a results-expectations chart emphasizes what is reasonable and thus what reasonable accountability expectations ought to be, and what needs to be measured. From this perspective, good performance expectations require a clear and concrete results-expectations chart. As discussed later, a results-expectations chart provides a practical basis for demonstrating just what influence the program has indeed had.

SETTING CLEAR AND CONCRETE EXPECTATIONS

There can be resistance in organizations to setting out clear and concrete performance expectations. Sometimes this is just a reluctance to have performance measured and compared with expectations. This is the cultural barrier that most organizations have to overcome if they wish to move toward managing for results (Auditor General of Canada, 1997, paragraphs 11.41–11.66). In other cases, the reluctance is due to not knowing what the level of results should be. This can be due to difficulties in measuring the outcomes in question as well as considerable uncertainty about how the activities of the program influence the expected outcomes. Clear and concrete expectations are needed to determine and demonstrate what has been accomplished.

Two Concepts of Expectations

There are two types of performance expectations. More traditionally, expectations are seen as *predictions* (targets) of what is hoped for in the future given the resources available, with the understand-

ing that they will be met most of the time. Alternatively, expectations can be seen as *challenges* to rise to. These are often referred to as stretch targets, set to provide clear direction on where the program is striving to be in the future. Stretch targets are set at a high enough level that they cannot easily be met with existing resources, and thus it is understood that they often will not be met. Their purpose is to set results to be strived for rather than results to be met with reasonable certainty. There is an expectation that in stretching, innovation will be encouraged and opportunities to learn and improve will be sought.

Performance expectations set as predictions (targets) can play a positive role, but they also have a number of shortcomings:

- Performance reporting can be as much a report on how good a prediction was as on how well a program is performing.
- Readers of a performance report, and many staff, still do not really know what is being reported: Are these stretch targets or safe predictions? Are they supposed to be easily met?
- Uncertainty about achieving outcomes, given the various factors not under the program's control, can result in setting targets only or mainly for outputs.
- Predictive targets are set with a view to reporting zero variance. This is not really reporting on how well the program is performing; the question remains — could the program have done better? It is also less likely to encourage innovation and learning; a manager may wonder — since the target has been met, why try harder?

The predictor model of setting expectations probably works better for output-based reporting than outcome-based reporting, since considerably more certainty can be established about the levels and quality of outputs produced.

I would argue that reporting on outcomes as part of performance reporting ought to mean using expectations as challenges — that is, stretch targets. This type of expectation reflects the learning model underlying the concept of managing for results. One of the many advantages of using stretch targets as expectations is the recognition upfront that many of these expectations will probably not be met, but that much learning will have taken place. This learning should be reported as achieved performance.

In between predictive and stretch targets might be predictive targets that are increased each year and the planned increases in levels that are set out beforehand publicly, so as to stretch an organization. Also in between would be targets that are quite challenging but, with enough effort, can be met.

Left on their own, most organizations would probably opt for the safer predictive targets to avoid having to explain why (stretch) targets have not been met. Challenging targets — as in the case of many of the Public Service Agreement targets in the U.K. (HM Treasury, 2002) — are usually set as the result of pressure on departments from a central budget office. The Auditor General (2000a) in Canada has pointed to the limited role played by the Treasury Board in requiring departments to set clear and concrete, much less challenging, targets.

Setting Expectations is Evolutionary

An essential element of the managing for results framework is the need to review and adjust measures and expectations over time as more experience and understanding are acquired. It is unrealistic to expect that at the outset one will be able to identify the perfect set of measures and corresponding performance expectations, and then set out to implement a performance measurement system that will last over time. The process is often evolutionary and advances through a lot of trial and error. Furthermore, the environment within which a program operates is constantly changing, and thus ongoing planning and consequent revisions to performance measures and expectations are needed. Clear and concrete expectations evolve over time from a better understanding of how to achieve the policy objectives.

The implication is that a results-expectations chart should be seen as an evolving construct. Over time, the chart becomes firmer with stronger and better understood logical links based on evidence, acquires stronger, more meaningful measures of key results, and develops more concrete expectations.

This evolution should occur in a deliberate manner, rather than as random trial and error. There should be a very visible built-in review and adjustment mechanism that identifies the “fittest” measures and expectations — that is, those that turn out to be useful to the organization for managing and reporting. This again reinforces the impor-

tance of deliberate learning based on past experience, the hallmark of managing for results (Barrados & Mayne, in press). Reporting simply on the gap between expectations and actual performance encourages a focus on meeting targets rather than learning.

If we put this approach together with the use of challenge-expectations, then reporting on performance comes naturally, with a focus on reporting the learning that is occurring rather than on variances from predictions. It also means that good performance reporting should include a description of how the results-expectations chart and its components are expected to evolve, how current expectations have been set, and what is hoped to be achieved with them as experience is gained.

Outputs Still Matter

The focus here is on setting outcome expectations and using them in planning, managing, and reporting. But in no way is it suggested that outputs are not important. Setting output expectations, measuring progress toward them, and tracking the expenditures associated with those outputs are essential for day-to-day management and for understanding the link between activities and outputs, which are what program managers can control. Although managers cannot track outcomes on a daily basis, they need to keep track of their day-to-day activities and the resulting outputs and expenditures. Managing for outcomes means that from time to time managers need to have information on how well their results chain reflects reality and thus be able to modify their activities and outputs as required.

Strategies for Developing Expectations

In keeping with the idea of an evolving results-expectations chart, several approaches can be adopted to move toward the goal of clear and concrete expectations:

- Identify benchmarks from other programs or jurisdictions.
- Measure performance for a period to establish a baseline.
- Base expectations on past performance.
- Set directional expectations first, measure progress, and engage in a discussion with interested parties about what level of performance is reasonable.
- Use qualitative approaches to measuring achievements.

- Consult with stakeholders (customers, budget offices, and legislators) on reasonable expectations.

Good performance reporting would report on these strategies for developing expectations and explain where the expectations came from and how they were determined.

Characteristics for Good Performance Expectations

At a minimum, stated expectations must be clear (well-defined and not ambiguous). The reader should be able to tell what aspect of performance is being reported.

Where possible, expectations should be concrete. However, they may vary as to the type or extent of concreteness. The ideal concrete expectation allows one to know when the expectation has been attained and has three elements:

- The target group who is to benefit from the results is clearly set out.
- The extent of the results to be achieved is specified.
- The time frame for achieving the results is stated.

Targets can be either both numerical with time frames or non-numerical. A variety of targets that are non-numerical or not a single number yet still concrete are possible, including the following:

- setting a range of expected performance;
- continuing a trend of past performance;
- establishing 0-1 type targets where an event or milestone is to be accomplished by a target date;
- setting a verbally defined scale such as poor, fair, or good; and
- making a significant contribution to an outcome, where “significant” is well-defined.

For a number of reasons, setting out a concrete target at the outset might not be feasible or useful. In moving toward this ideal, one might adopt a number of approaches, such as setting out a concrete expectation without a time frame, or setting out a clear direction of change expected, but without the end point specified. In both cases, the expectation could be made more concrete once more experience

is gained. The reasons for using either of these approaches should be reported.

Where expectations are clear but not very concrete, one approach would be to assess the achievement of the expectations using an external expert review panel with a qualitative scale. The National Science Foundation (2000) in the U.S. has used this approach in its performance reporting. Each of its programs is reviewed every three years by groups of independent external experts for their success in meeting targets. They use a two-point rating scale: “successful” and “minimally effective.” Table 3 illustrates the results for the goals and targets of one program.

The credibility of this approach clearly depends on the credibility of the experts used. But it is one way of dealing with the problem of expectations that are hard to quantify. Over time, there may be agreement on what constitutes success in these areas.

Table 3
Reporting Against Non-Concrete Expectations: The National Science Foundation (NSF)

Outcome Goal 1

Discoveries at and across the frontier of science and engineering

Targets:

- Make important discoveries; uncover new knowledge and techniques, both expected and unexpected, within and across traditional boundaries
- Forge new high-potential links across those boundaries

Assessment: Successful. All groups of experts rated NSF successful.

Tangible examples: Specific advances in biology, funding the Nobel Prize winner in Chemistry, research in the Antarctic and Arctic, and discoveries in how the young learn.

Outcome Goal 2:

Connections between discoveries and their use in service to society

Targets:

- The results of NSF awards are rapidly and readily available
- The results are fed as appropriate into education or policy development
- The results are used by other federal agencies or the private sector

Assessment: Successful. 42 of 43 expert groups rated NSF successful.

Tangible examples: Predicting storms, oceanographic research in service of fisheries management, practical application of digital library, sustainability projects, award-winning NSF supported children's science television shows.

Source: National Science Foundation (2000).

Selecting Performance Measures

Key to setting clear and concrete performance expectations is selecting good performance measures to represent those expectations. Good measures produce measurements that are accurate, reliable, and valid.

Accurate measurements are those that correctly represent the reality measured. Reliability implies that if others use the same measures, they will produce the same results. Valid measures measure the performance characteristic of interest rather than something else. For example, an IQ test may be an accurate and reliable measure, but as a measure of intelligence it has poor validity. Similarly, the time taken to go from 0 to 100 kilometres per hour is a valid measure of the acceleration of a car, but not a valid measure of the performance of that car.

Often, good measures that provide useful information on performance are not evident until after measurement has been underway for some time. Reviewing and revising the usefulness of the measures is good practice.

For most programs, it is easy to identify quite a large number of performance measures from a results chain and correspondingly to set a large number of expectations. This is normally not useful. One needs a manageable number of expectations, perhaps four or five for a particular span of management control. Too many measures and expectations cannot be managed and will likely end up just feeding information systems.

“Gaming” Expectations

Setting expectations with the knowledge that they will be met is one way of “playing games,” not really of managing for results. In presenting the practice of developing expectations over time, another obvious game could be played: continue to change expectations every year so that over time, progress cannot be assessed. This is not being suggested; rather, there should be a deliberate strategy to improve expectations, with many of them becoming quite stable over time (unless the underlying activities change). The initial evolution of expectations should be over a period of several years, not as long as a decade.

A Final Word

Although the discussion has been on setting expectations for a specific result to be accomplished, it is important to recall that this is all being done within the bigger picture of the results-expectations chain. From that perspective, it might be appropriate for some of the expectations to be quasi-concrete — perhaps for the foreseeable future — as long as others were moving toward being concrete. Further, the number of concrete outcome-expectations needed will depend to some extent on the strength of the logic chain for the program. The stronger the evidence is that the program logic works, the less may be the need for concrete, higher-level outcomes. Thus, for example, an anti-smoking program has as its ultimate outcome the improved health of Canadians. On the surface, this is an outcome quite far removed from the activities of such programs. However, the link between reduced smoking and improved health has been strongly established. It would not be reasonable for the anti-smoking program to try to measure how reduced smoking is affecting health. In telling its performance story, demonstrating reduction in smoking would count for success.

TELLING A PERFORMANCE STORY

Good reporting on performance requires reporting on what was achieved in relation to what was expected. Reporting on outcomes therefore involves reporting on what was achieved in relation to the results-expectations chart set out. It involves more than simply reporting against several specific performance measures or targets previously set out. It involves telling a credible *performance story* about the results-expectations chart — that is, presenting evidence on the extent to which the results-expectations chart in fact reflects reality. Of course, if in addition, unintended results are observed, these too should be reported. Indeed, it is good practice to ensure that the strategy for measuring program results includes room for addressing unintended results associated with the program.

Presenting the Story

There are a variety of ways to present a performance story. All involve a mix of quantitative evidence that certain outputs and outcomes have occurred as well as narrative discussion and further evidence of the contributions made at various points along the re-

sults chain, all described within some context. A performance story sets out to convince a skeptical reader that the activities undertaken by the program have indeed made a difference — that the expectations chain has, at least to some extent, been realized, along with any significant unintended results.

Table 4 identifies the main elements of a performance story and illustrates further detail on each element. These elements could be used to present the story in a narrative form, with accompanying evidence on the specific results being reported and with reference to the expectations chart. One example of this approach can be found in the performance report of the Office of the Auditor General (2002).

Reporting on the Contribution Made

One of the more challenging aspects of reporting on performance is to credibly report on the contribution the program has made to the expected outcomes. The aim is to show that the program has made

Table 4
Elements of a Performance Story

A good performance story covers the following elements:

What is the context?

- the overall setting of the program (description, objectives, resources)
- the results chain (program theory)
- the risks faced

What was expected to be accomplished at what cost?

- statement of the (clear and concrete) outputs and outcomes expected
- planned spending

What was accomplished in light of these expectations?

- the relevant outputs delivered at what cost
- the outcomes realized related to the expectations
- a discussion of the evidence available demonstrating the contribution made by the program to those outcomes

What was learned and what will be done next?

- a discussion of what will be done differently as a result of what was achieved

What was done to assure quality data?

- a description of what the organization does to ensure the quality of the data and information reported

The main story line of a performance story is how well the program has performed in relation to what was expected and what will now be done differently to better ensure future performance.

a difference. Practical approaches to undertaking a contribution analysis have been discussed elsewhere (Mayne, 2001), in particular, without the availability of a robust evaluation that addressed the contribution issue. In this case, the aim is not to definitively prove that the program has made a difference but to build over time a convincing case of plausible association between the results observed and the activities and outputs of the program, as more and more evidence is gathered. Presenting a case that is reasonably accepted is a more realistic aim.

The key component is the results-expectations chart, especially its logic. The logic model sets out what is supposed to happen as a result of the program activities. If this logic is not convincing, then the likelihood of a convincing contribution story is low.

While essential, the logic story only says what is supposed to happen — why one believes that the program is supposed to make a contribution. The real question is, has it? A variety of data and information could, and indeed should, be collected to build the evidence (Mayne, 2001), such as:

- The expected sequence of events outlined in the logic story of the program have indeed occurred.
- Even more detailed logic models of key aspects of the program theory are also confirmed by events.
- Experts in the area agree that the program was a major factor in contributing to the observed outcome.
- Other factors suggested as having an influence on the observed outcome are not as plausible as the factors outlined in the program's theory.

The example in Table 3 uses groups of outside experts to conclude that the program of the National Science Foundation has made a contribution. Does it present undisputable evidence? Probably not, but quite a strong and reasonable case is made.

It is likely that no single piece of evidence gathered will on its own be enough to build a credible case concerning a result achieved or a contribution made by a program. As discussed elsewhere, “Although no one piece of evidence may be very convincing, a larger set of different and complementary evidence can become quite convincing” (Mayne, 2001, p. 20). It is the totality of the evidence gathered — some of it strong, some perhaps rather weak — that builds a credible perform-

ance story. The measurement strategy used ought to anticipate the need for multiple sources of data, information, and analysis.

Ways of Telling a Performance Story

The most common way a performance story is told is probably through a narrative describing the performance accomplished in light of what was expected. If the results chain for the program is used, the story can be expressed as discussing the extent to which the results chain actually reflects reality, giving structure and logic to the story.

One alternative approach to telling key elements of a performance story would be to prepare a *chart of performance accomplishments* or a *performance story chart* paralleling the expectations chart. Table A.1 in the Appendix presents a generic performance chart. The left-hand column presents the basic results chain against which the right-hand column is reporting. This is needed to remind the reader of the basic theory and assumption underlying the program. The right-hand side — the substantive performance story — can cover a number of elements:

- the general rationale for the program;
- the external context within which the program operates;
- a historical picture of performance;
- last year's (or the most recent) statements of specific performance expectations;
- recent accomplishments; and
- a discussion of the linkages between various levels in the results chain.

There is even room in the performance chart for anecdotal evidence. Anecdotal evidence on its own is usually not very credible. It may be quite selective and not representative of the true picture. However, presented in a performance chart, the anecdote has context and serves to illustrate that at least some cases are conforming to the expectations chart. The anecdote is anchored somewhere.

Table A.1 also illustrates how the learning that has occurred can be reported, by setting out a summary of the implications drawn from past performance. Table A.2 in the Appendix presents a performance chart for the results-expectations chart shown in Table 2. There is not much experience to date with using performance charts to tell

performance stories. Table A.2 illustrates one approach that might be used to present parts of the story. The aim is to find ways to succinctly present a meaningful story.

WHAT SHOULD PARLIAMENT AND THE PUBLIC EXPECT?

Setting out clear and concrete performance expectations and reporting progress against them is essential for good performance reporting to Parliament and the public. Without good statements of expectations, it is not clear whether the results achieved represent good performance. Several things are required to set out clear, concrete, and credible expectations. Good reporting should:

- be clear about which type of expectation is being set out. Is it a stretch target or a prediction?
- describe how the expectations were arrived at. Why are the expectations a reasonably expected level of performance?
- describe how the expectations relate to the objectives and mandate of the program. Why are they important to achieve? How is the organization expected to contribute to the expectations? A results-expectations chart might be one of the best ways to clarify the contribution made succinctly.
- set out the expectations in clear and concrete terms.

In a results-expectations chart, a good statement of performance expectations would articulate the clear and concrete outcomes sought, the concrete outputs to be produced, and the contribution logic (why the outputs are expected to contribute to the outcome).

Finally, good reporting to Parliament and the public ought to set out credible information on each of the elements of a performance story, as outlined in Table 4.

CONCLUSION

Managing for and reporting on outcomes instead of outputs requires new approaches to setting performance expectations and telling performance stories. A number of new ways to approach this problem have been suggested, including:

- setting expectations in the context of a results-expectations chart rather than in terms of individual measures;

- recognizing that the results-expectations chart will and should evolve and become more robust over time;
- recognizing additional ways, other than using single numbers, of setting out concrete expectations;
- distinguishing between challenging and predictive expectations, with a focus on challenging or stretch targets for outcomes; and
- reporting performance as a structured story in the context of a results-expectations chart, either as a narrative or in the form of a performance story chart.

At the basis of these suggestions is the results chain, a key tool, it is argued, for setting performance expectations and telling credible performance stories.

NOTE

1. An exception might be cases of programs that are experimental in nature—usually pilots—where the intervention is expected to be beneficial but there is perhaps little experience to date, and hence concrete expectations may not be practical or are expected to emerge as an understanding of the program’s contribution is gained. Over time, realistic performance expectations could be developed.

REFERENCES

- Auditor General of Canada. (1997, October). Moving toward managing for results. *Report to the House of Commons*, Chapter 11. Ottawa: Author.
- Auditor General of Canada. (2000a, December). Managing departments for results and managing horizontal issues for results. *Report to the House of Commons*, Chapter 20. Ottawa: Author.
- Auditor General of Canada. (2000b, December). Reporting performance to Parliament: Progress too slow. *Report to the House of Commons*, Chapter 19. Ottawa: Author.
- Auditor General of Canada. (2002, December). Modernizing accountability in the public sector. *Report to the House of Commons*, Chapter 9. Ottawa: Author.
- Barrados, M., & Mayne, J. (in press). Can public sector organizations learn? *OECD Journal on Budgeting*.

Economist (2001, April 26). The trouble with targets.

General Accounting Office. (2003). *Using GPRA to address 21st century challenges*. GAO-03-1166T. Washington, DC: Author. Retrieved November 1, 2003 from <<http://www.gao.gov/>>.

HM Treasury. (2002). *2002 spending review, public service agreements 2003–2006*. Cm 5572. London. Retrieved October 30, 2003 from <<http://www.hm-treasury.gov.uk/>>.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16(1), 1–24.

Millar, A., Simeone, R., & Carnevale, J. (2001). Logic models: A systems tool for performance management. *Evaluation and Program Planning*, 24, 73–81.

Montague, S. (2000). Focusing on inputs, outputs, and outcomes: Are international approaches to performance management really so different? *Canadian Journal of Program Evaluation*, 15(1), 139–148.

National Science Foundation. (2000). *Accountability report FY 1999*. Washington, DC: Author.

Perrin, B. (2002). *Implementing the vision: Addressing challenges to results-focused management and budgeting*. OECD. Paris. Retrieved October 30, 2003 from <<http://www.oecd.org/dataoecd/4/10/2497163.pdf>>.

Office of the Auditor General of Canada. (2002). *Performance report for the period ending 31 March 2002*. Ottawa: Author.

Rossi, P., Lipsey, M., & Freeman, H. (2004). *Evaluation: A systemic approach* (7th ed.). Thousand Oaks, CA: Sage.

Thomas, P. (in press). *Performance measurement, reporting and accountability: Recent trends and future directions*. Regina, SK: Saskatchewan Institute of Public Policy. <<http://www.uregina.ca/sipp/>>.

Wholey, J. (1997). Clarifying goals, reporting results. In D.J. Rog (Ed.), *Progress and future directions in evaluation: Perspectives on theory, practice and methods* (pp. 95–105). New Directions for Evaluation, No. 76. San Francisco: Jossey-Bass.

Table A.1
A Chart of Performance Accomplishments

| | <i>Results Chain for Program</i> | <i>The Performance Story</i> |
|--|--|---|
| Rationale | A short statement of the problem or situation being addressed and the overall rationale for the program and how it is to address the concerns. The discussion would provide the general objectives for the program, i.e., what it is expected the impact of the program will be. | |
| Activities/ Outputs | <ul style="list-style-type: none"> • What is the context? • What are the activities of this program? • What outputs are produced at what cost? | <p>What outputs have actually been delivered? In many cases, one would want to present key outputs over the last several years, as well as those delivered in the most recent year. For the recent fiscal year, state the expectations for the year and accomplishments against them. If a longer explanation is required, use additional notes or figures at the end of the chart.</p> <p>Prior years</p> <ul style="list-style-type: none"> • key outputs in past years (either as part of a time series — e.g., add a Figure 1 below — or as part of the multi-year story) <p>FYXX Key expected outputs</p> <ul style="list-style-type: none"> • expected output1 (from Expectations Chart) • expected output2 (from Expectations Chart) <p>Achievements</p> <ul style="list-style-type: none"> • output1 actuals • output2 actuals • planned expenditures versus actuals • Any other significant activities undertaken or outputs produced |
| Actual Spending (fiscal year) \$xx.x million | | |
| | ↓ | ↓ |
| Target Group | <ul style="list-style-type: none"> • Who are you trying to influence? | Which target groups have you affected and how? |
| | ↓ | ↓ |
| Partners | <ul style="list-style-type: none"> • Who are your partners? | How have your partners affected the desired results you are seeking? |
| | ↓ | ↓ |
| Immediate Outcomes | <ul style="list-style-type: none"> • What do you expect to be the immediate results of your activities/outputs? • What do you expect will happen as a result of producing your product? | <p>Previous years' significant immediate outcomes which are essential to the performance story.</p> <p>Key expected immediate outcomes</p> <ul style="list-style-type: none"> • immediate outcome1 <p>Achievements</p> <ul style="list-style-type: none"> • immediate outcome1 actual |

Table A.1 (continued)

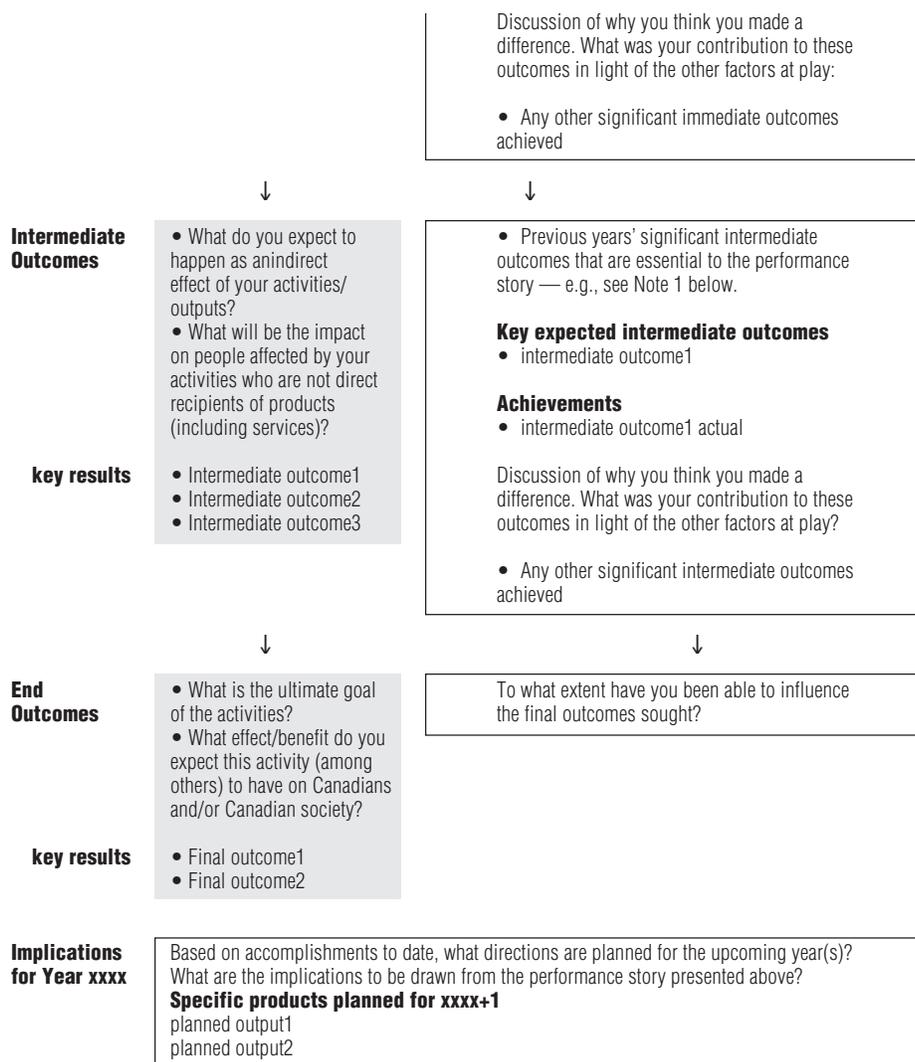


Figure A.1
Time Series of Outputs Produced
 (A time series chart could be provided here.)

Note

1. A presentation of the significant intermediate outcomes achieved over the past several years and a discussion of how they will lead to the final outcomes could be provided here.

Table A.2
A Performance Story Chart

Audit Office Activities to Enhance Accountability Practices and Managing for Results in Government

| | <i>Results Chain</i> | <i>The Performance Story as at 31 March 2002</i> |
|---------------------------|--|---|
| Rationale | The Office has a long-standing interest in improving accountability practices in government, especially accountability to Parliament. Through pertinent audit reports and related communication on accountability and governance issues, the Office expects to draw attention to significant accountability issues and encourage better practices. | |
| Activities/Outputs | <ul style="list-style-type: none"> • Developing frameworks (principles and audit criteria) for good accountability • Identifying weaknesses and strengths in organizations • Clearly articulating accountability for results communicated in audit office reports, presentations, speeches, and advice | <p>1997 Audit on Managing for Results tabled: Framework and good practices presented.</p> <p>1998 Discussion paper on <i>Modernizing Accountability Practices</i>: New concepts for accountability, accountability for results, and shared accountability presented.</p> <p>1999 Two audits on collaborative arrangements and new governance arrangements dealt extensively with accountability issues.</p> <p>2000-01 Audit on managing for results in five departments</p> <p>Key specific expected products in 2001-02</p> <ol style="list-style-type: none"> 1. Audit published on accountability in new governance arrangements 2. Office Guide on auditing accountability <p>Achievements</p> <ol style="list-style-type: none"> 1. Audit on new governance arrangements that set out a clearer framework for accountability to Parliament 2. Draft Guide on Auditing Accountability hat was further developed but not completed <ul style="list-style-type: none"> • Modernized accountability concepts from the discussion paper were used in several presentations and speeches |
| Target Group | Office auditors, entities, central agencies and parliamentarians | This year there was continued communication with the target groups through speeches, presentations, and participation in seminars. |

Table A.2 (continued)

| | | |
|--|--|--|
| <p>Immediate Outcomes</p> | <ul style="list-style-type: none"> • Entity audits examining accountability with consistent criteria and recommendations • Better understanding and acceptance of accountability for results • Better understanding and support for managing for results in government | <ul style="list-style-type: none"> • The Treasury Board Secretariat continued to support the Office accountability framework • Office concepts of accountability are being used in some World Bank and OECD documents • The idea of managing for results is becoming widely accepted and being given priority by the Treasury Board <p>Key expected immediate outcomes</p> <ol style="list-style-type: none"> 1. Greater use of the Office accountability framework in audits dealing with accountability 2. Greater acceptance of our accountability concepts 3. Better accountability frameworks in departments <p>Achievements</p> <ol style="list-style-type: none"> 1. Continued improvement in Office consistency on accountability, with more audit teams consulting on accountability 2. Continued growth in use of our accountability concepts, e.g., shared accountability being seen as a means to address management of horizontal issues 3. Some evidence of better frameworks, as shown in the new governance audit |
| | ↓ | ↓ |
| <p>Intermediate Outcomes</p> | <ul style="list-style-type: none"> • Transparency in accountability relationships in government • Increasing practice of managing for results across government | <p>Implementation of managing for results in government is proceeding too slowly. Use of accountability concepts in government is not widespread.</p> |
| | ↓ | ↓ |
| <p>End Outcomes</p> | <ul style="list-style-type: none"> • Enhanced accountability for results | |
| <p>Implications for 2002–03</p> | <p>Greater effort is needed on communicating the principles and good accountability practices identified by the Office.</p> <p>Specific products planned for 2002–03</p> <p>Study on accountability published Discussion paper on strengthening parliamentary scrutiny Guide on auditing accountability</p> | |

John Mayne worked at the Treasury Board Secretariat in the 1980s and 1990s and was instrumental in the development of the federal government's approach to evaluating the performance of programs. In 1995, he joined the Office of the Auditor General where he led efforts at developing effective performance measurement, managing for results, and performance reporting practices in the government of Canada, as well as methodologies for providing audit assurance on performance reports. In recent years, he led the Office's audit efforts in accountability and governance. He has authored numerous articles and reports, and edited five books in the areas of program evaluation, public administration, and performance monitoring. In both 1989 and 1995, he was awarded the Canadian Evaluation Society *Award for Contribution to Evaluation in Canada*.