



**IMPLEMENTING THE VISION: ADDRESSING CHALLENGES TO RESULTS-FOCUSED
MANAGEMENT AND BUDGETING
BY
BURT PERRIN¹**

1. Introduction

1.1. Background

1. Over about the last decade, almost all OECD member countries have undertaken public sector reform activities in order to improve program effectiveness and efficiency and the quality of management and the budgeting processes. One of the key objectives of most reform activities has been to provide for an increased focus from inputs to the actual results and impacts of government activities.

2. Approaches to public sector reform have varied across countries, reflecting differing contexts, cultures and capacities. Nevertheless there have been considerable similarities in terms of the overall thinking behind the need for public sector reforms, as well as with respect to many of the approaches that have been attempted in order to build more of a results orientation into management and budgeting processes. As approaches to reform have been maturing, there also appear to be considerable similarities with respect to issues and challenges that have been encountered to date. In spite of (or perhaps because of) the claims of public sector reform, there has been surprisingly little evaluation to assess its appropriateness or impact. There is an increasing number of critical commentaries calling into question many of the basic tenets of some results-oriented approaches.

3. This context provides a good opportunity for the sharing of experiences to date across OECD countries, and in particular to identify implementation problems and challenges that have been experienced and to discuss how these might be addressed. Accordingly, the Directorate for Public Governance and Territorial Development (GOV) of the OECD convened an expert meeting on 11-12 February 2002 at its Paris headquarters. This meeting generated considerable interest. Some 27 member countries sent representatives. About half prepared papers and made presentations on the situation in their own country.

4. Participants at the meeting reaffirmed the commitment of their countries and their own strong belief in the importance of a results-focused approach. At the same time, there were open and frank discussions about many of the challenges that have been encountered. There are some significant differences in the status of reforms from one country to another, as well as with respect to philosophy and how these have been implemented. Nevertheless, there was remarkable consistency across countries with respect to many of the implementation approaches, with respect to the difficulties that were encountered, and ideas about how these could be addressed. It is fair to say that the meeting engendered more discussion of challenges than of solutions.

1.2. Purpose and organisation of this paper

5. This paper represents an expansion of a response to the country presentations that was presented towards the end of the expert meeting. Its primary purpose is to identify and summarise major challenges to a results-focused approach, and to identify potential solutions to these challenges. As much as possible, it aims to be constructive, in suggesting ways in which the challenges that have been identified by participants at the expert meeting *can* be addressed. While some of the information in this paper may be useful to those involved in the actual implementation of results-focused approaches, its primary emphasis is on what can be done at the strategic and systems levels of government.

6. This paper is based largely upon the country papers, presentations and discussions at the expert meeting. In order to provide for a further understanding of some of the problems that have emerged in OECD member countries and to identify potential solutions, it also goes beyond to the literature.

1.2.1. Key Themes

7. The balance of this paper is organised around the following key themes that emerged in the discussions at the expert meeting

A focus on outcomes is very important, but it is also very difficult. (Section 2)

8. Many countries say that this has proved far more difficult than expected. In particular, they find it hard to see how one can connect what programs and managers do with distant and complex outcomes. For the above and other reasons, many of the countries acknowledged that outcomes receive only minimal attention. Instead, most efforts at performance monitoring thus far have focused mainly on inputs, activities and outputs. Yet there are ways in which one can assess outcomes.

Creation of a results-oriented culture throughout government is essential for actual implementation of a results-oriented approach. (Section 3)

9. Unless managers and staff throughout government actually buy into a results focus, they most likely will just go through the motions, and performance management risks remaining a paper exercise. There was strong agreement about the importance of this issue, but there was also recognition that it is not clear how to bring about the requisite culture change.

Attention is needed to the question of data quality, to ensure that the performance data that is obtained is meaningful and valid. (Section 4)

10. This concern was acknowledged perhaps more implicitly than explicitly at the meeting. Inaccurate or misleading data are useless — or worse — for decision-making, and can even distort program activities. In order to ensure the meaningfulness of performance data, one needs to be sensitive to potential difficulties and how these can be alleviated or addressed.

Leadership and support from the top levels of government is needed in order to bring about a results-focused approach. (Section 5)

11. Leadership can take various forms, including: top-level support for a results-focused approach, capacity building, and effective communications. As well, proponents of results-focused management and budgeting need to lead by example, in particular by undertaking independent evaluation of reform initiatives. Audit also has a role to play.

A strategic rather than a piecemeal approach is necessary. (Section 6)

12. For effective performance management, an integrated approach involving both monitoring and evaluation is needed. It is important to develop an approach that recognises the realities and complexities of working in the public milieu, with many interacting initiatives and factors. There are a variety of other important strategic considerations, such as: how a results-oriented approach can contribute to a whole-of-government approach, how it can be used when contracting externally for services, how performance information can be combined in a meaningful way with the budgeting process, and how it can be made relevant to Parliament.

2. An orientation to outcome

2.1. What do we mean by “outcome”? Why is it important?

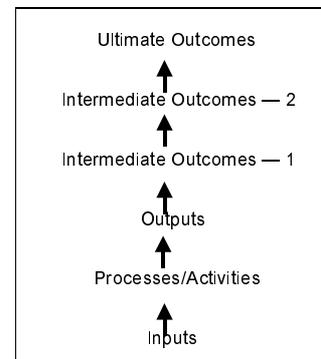
13. Until recently, the performance of programs, and of program managers, has been judged largely on inputs, in particular on how they have spent their allocated budgets. Traditional bureaucracies would devote most of their attention to their processes and activities, and — perhaps — to the outputs produced. They would work at trying to be more efficient at delivering what they traditionally have been producing, with limited consideration to the value or appropriateness of their activities and outputs.

14. This approach, however, has been called into question by public management reform in most OECD countries, which now places an emphasis on results. One of the major factors that has led to public management reform is a concern that too often, government is preoccupied with *process* and with following rules, and that it is not clear what *benefits* are actually arising from public service expenditures and activities. This has led to the creation of results-focused approaches that emphasise the *outcomes* of public expenditures, that is the benefits that arise from one’s actions rather than on processes or what one *does*. Monitoring and evaluation activities have been increasing at exponential rates, with attempts in some jurisdictions to tie management and budgeting to results.

15. Moving to a focus on outcomes involves a shift in thinking as much as in measurement. This requires managers at all levels to bear in mind that they are engaged in activities and producing outputs not for their own sake but in order to achieve “big picture” outcomes in line with the mission of the program. This requires thinking about outcomes on a regular basis, in order that the mix of activities and outputs are frequently reviewed and adjusted as necessary in order to maximise the impact of the program in addressing the needs that form the rationale for its existence. Otherwise, activities and outcomes easily can become irrelevant.

16. Kristensen² has indicated that: “Outcomes reflect the intended and unintended results from government actions and provide the rationale for government interventions.” Mayne³ has added: “That is what the public sector is all about — producing benefits, or results, for its citizens.”

17. Thus the *raison d'être* of a results-focused approach is to bring about a shift in focus along the results chain, as illustrated in the accompanying box, from inputs to outcomes. As John Mayne, in Canada's Office of the Auditor General put it: "The aim [of a results focus] is to change the culture of public administration from one that is rules focused to one focusing on the results that matter to citizens."⁴



18. Thus a focus on outcomes is central to any results-focused approach.

2.2. *Current status*

19. There appeared to be a general appreciation at the expert meeting of the need for an outcome focus. Yet at the current time, there seems to be just variable attention to this. For example, the country papers contained numerous references to terms reflecting a primary concern with input, or to a limited extent, output, such as: "efficiency", "productivity", "activities", "procedures", "expenditures" and "outputs". But in general, there was much less use of outcome-oriented terms, such as: "benefits", "value", "effectiveness", "outcomes", "impacts" or "quality of life".

20. A number of the papers acknowledged that they have not been successful in implementing an outcome perspective, as the following quotes illustrate:

Today, results-based management primarily focuses on output and workload indicators. ... While output is easy to quantify and monitor, [our] experience shows that relying solely on output involves a risk that agencies will lose sight of the effect their programs are intended to have on society.

There are few good examples of ministries and agencies that have achieved an effective linkage and correlation between input and output, which was an objective of implementing performance management as a tool in public governance.

These problems made it difficult for anyone without a detailed knowledge of [the department] to make an informed assessment of the department's overall performance. ... While there is considerable information regarding [outputs] (individual payments and services), it is not clear how they contributed to the achievement of the planned outcomes.

21. The countries giving papers and presentations at the meeting are at various stages along the results chain in terms of their primary approach. In some countries, such as Brazil, France, Germany, Ireland, the primary focus appears to be mainly (but not necessarily entirely) with respect to reform of the budgetary process. In other countries, including Denmark, Iceland, Netherlands, Norway and Sweden, there is considerable emphasis placed on the monitoring of activities and outputs. In contrast, Australia, the United States and the United Kingdom are actively engaged, at least to some extent, in both the

monitoring of outputs and the evaluation of outcomes. Finally, the European Union, Korea and Spain are looking to expand their capability for the evaluation of outcomes, without necessarily working through a process of output indicators.

22. Why has there been such limited attention to outcomes, in spite of their central position to the very concept of a results-focused approach to performance? One key reason is that as a number of countries acknowledged, this has proved immensely difficult, considerably more difficult than they had anticipated. Despite the implications of public sector reform regarding the importance of thinking and managing in outcome terms, traditional ways of thinking are persistent. Within governments and agencies, there frequently is resistance to outcome-focused approaches.

23. As for example one country put it: “Outcomes are immensely difficult to measure and monitor.” They are far more difficult to quantify, if it is even possible, than are activities or outputs. By their very nature, frequently they would not be expected to be manifested until some time after the program intervention, generally not in sync with the same budgeting cycle. And outcomes typically come about not just as the result of a single intervention by one program in isolation, but by the interaction of a number of different factors and interventions, both planned and unplanned.

24. Thus to summarise: outcomes are viewed as very difficult to identify, and certainly difficult to quantify; they are seen as remote in time and space from what the program does; and interact with other factors. For these and other reasons, it seems difficult or impossible to link outcomes directly to actions or with managerial decisions.

25. For these and other reasons, there is a tendency to fall back on what is easiest to measure and to count, *i.e.* outputs, even if they are less important than outcomes. The danger of this, however, is to negate the major purpose of results-based reform, which is to refocus efforts on what citizens and society ultimately gain from government. Without a focus of at least some form on outcomes, one can too can lose sight of the bigger picture, which is what the program is supposed to be about. And if as is commonly stated: “What gets measured gets done,” then performance measurement can result in perverse effects, with less rather than more focus on impact.

26. Indeed, the experience in many different jurisdictions over the decades is that without an outcome perspective, one invariably gets “stuck”. That is to say, one does not “move up the ladder” from a focus on outputs to outcomes unless there is a major reason to do so. For example, a recent major review by the New Zealand Treasury⁵ of the effectiveness of its public sector reforms indicated that the focus was on efficiency rather than on effectiveness, with no clear links to policy. With an almost exclusive emphasis on accountability for outputs, there has been little attention to outcomes and no indication that the reform regime has actually resulted in benefits to New Zealand society and its citizens.

27. Thus if there is one key message emerging from the expert meeting, it is that there is a need to a much greater focus on outcomes, along with the development of the necessary supporting systems, structures and availability of expertise.

28. Besides, there is substantial evidence from a variety of sources that despite the perceived difficulties, outcome assessment in fact *can* be done, albeit at times with a somewhat different mindset and approach than from the direct measurement of inputs and outputs. This is briefly discussed in the following section.

2.3. *How to implement performance management with a true outcome focus*

29. Mayne has suggested using contribution analysis as one approach by which one can assess outcomes. He has summarised this approach as follows:⁶

Step 1: Develop the results chain	Describe the program theory model/program logic/results chain describing how the program is supposed to work. Identify as well the main external factors at play that might account for the outcomes observed. This program theory should lead to a plausible association between the activities of the program and the outcomes sought. Some links in the results chain will be fairly well understood or accepted. Others will be less well understood or subject to explanations other than that the program was the “cause.” In this way you acknowledge that attribution is indeed a problem.
Step 2: Assess the existing evidence on results	The results chain should provide a good idea of which intended results (outputs, intermediate and end outcomes) could be measured. What evidence (information from performance measures and evaluations) is currently available about the occurrence of these various results? The links in the results chain also need to be assessed. Which are strong (good evidence available, strong logic, or wide acceptance) and which are weak (little evidence available, weak logic, or little agreement among stakeholders)?
Step 3: Assess the alternative explanations	Outcomes by definition are influenced not only by the action of the program but also by external factors — other programs, as well as social and economic factors. In addition to assessing the existing evidence on results, there is a need to explicitly consider the extent of influence these external factors might have. Evidence or logical argument might suggest that some have only a small influence and that others may have a more significant influence on the intended results.
Step 4: Assemble the performance story	With this information, you will be able to set out your performance story of why it is reasonable to assume that the actions of the program have contributed (in some fashion, which you may want to try and characterize) to the observed outcomes. How credible is the story? Do reasonable people agree with the story? Does the pattern of results observed validate the results chain? Where are the main weaknesses in the story? There always will be weaknesses. These point to where additional data or information would be useful.
Step 5: Seek out additional evidence	To improve your performance story you will need additional evidence. This could involve information on both the extent of occurrence of specific results in the results chain and the strength of certain links in the chain. A number of strengthening techniques that you might be able to adopt are outlined in this work.
Step 6: Revise and strengthen the performance story	With the new evidence, you should be able to build a more credible story, one that a reasonable person will be more likely to agree with. It will probably not be foolproof, but will be stronger and more credible.

30. The Government of Canada is now promoting the use of performance measures as a way for programs to tell their “performance story”. As explained in an OECD publication:

Putting this approach into practice means thinking through the ‘results chain’: the logical connections between inputs, activities, outputs and outcomes. The development of this logic means agreeing on expected outcomes, measuring progress toward them, using this information in decision-making, and reporting results. The government believes that encouraging federal departments and agencies to use this results chain to express their results

allows for maximum flexibility in telling their performance story. In order to assist federal organizations to think through this results chain and to articulate their performance story more cogently, we are currently creating, in consultation with departments and external stakeholders, a lexicon of key 'results-based management' terms.⁷

31. Thus while it is possible to assess outcomes, there are some differences from how one goes about identifying inputs and direct outputs.

32. **Use of evaluation as well as performance monitoring.** As Mayne has indicated, “a good overall measurement strategy would include both ongoing performance measurement and periodic evaluation.”⁸ Indeed, identifying outcomes and linking them in some way to program interventions represents the main task of evaluation. There is a wide-range of potential evaluation methods and approaches that can be applied, depending upon the particular questions and situation and the degree of rigour required. There is an extensive literature on evaluation, along with professional associations now active in most OECD countries.

33. Evaluation approaches become increasingly important as one moves up the results chain, because of the interaction of a variety of different factors and the increasing complexity of outcomes and the identification of causality. The strategic implications of a combined monitoring and evaluation approach are considered in Section 6.1.

34. **Theory-based approach.** The approach outlined above is one example of what is referred to as a theory-based approach to evaluation.⁹ This currently is one of the most common approaches within the international evaluation community. For example, it forms part of the theme of the forthcoming 2002 European Evaluation Society conference.

35. Part and parcel of a theory-based approach is the articulation of a program logic (or intervention) model. A logic model identifies the presumed causal links and connections between and among activities, outputs and outcomes at various levels. It illustrates how the program is expected to work. It is an invaluable tool for program planning, and also can help direct appropriate monitoring and evaluation approaches at various stages in the program development. The results chain model on Page 4 represents a very simplified version of a program logic model. In fact, there is always a number of different intermediate outcomes at various levels. Progression is rarely linear as illustrated on the model. A logic model should be developed for every program or intervention, also indicating external factors that also may be expected to have some impact on outcomes at various levels.

36. **Acceptance of uncertainty.** As Mayne indicates, measurement has limits, especially at the outcome level, and we must be prepared: “to accept some uncertainty, where the aim of measurement should be to acquire some insight and develop some assurance that the program is actually having an impact.”¹⁰ As the European Union guide to evaluation¹¹ indicates, it is better to have imprecise answers to the important questions than to have precise answers to those that are unimportant.

37. **Constantly keep the “big picture” in mind.** Use of a logic model or results chain can help maintain thinking and perspective on outcomes — the *raison d'être* of government activities — at all times, even when one is measuring performance at a more intermediate level and when it may be premature to assess outcomes further along the results chain. This *perhaps is the most important factor* and the essence of a results-focused approach. In this way, one can avoid viewing outcomes as all or nothing, which appears to be one reason why many people have backed away from an outcomes approach.

38. One should consider one's progress along the results chain on a periodic basis. This should involve checking initial assumptions of the "logic" of the results chain and how the program outputs are expected to contribute to desired outcomes. This inevitably will result in some adjustments. This is one way to help keep performance measures meaningful.

39. Thus it *is* possible to assess progress towards outcomes, even if it is not always easy or possible to specify in exact terms.

40. **Maintain a balanced approach.** It is of course necessary to monitor inputs and outcomes, as well as to assess progress towards outcomes. Each jurisdiction needs to assess the appropriate balance, which may very well vary from agency to agency and from program to program, depending upon the status of its life cycle, nature of the program activity and other factors. But given that the current emphasis in most OECD countries appears to be on inputs and perhaps on outputs, it would appear that in most cases more attention to outcomes would be appropriate. As a recent OECD report states:

*Anticipating, learning from and communicating the outcomes of policies will be at the core of the transformation of public organisations into learning organisations. The systematic focus on outcomes is in its infancy and the establishing of the right form and level of evaluation still represents an important challenge.*¹²

41. Without some focus on outcomes from time to time, inputs and outputs run the risk of becoming disconnected from their reason for existence. And as noted earlier, this perception is one of the driving forces that has been behind public sector reform.

3. How to provide for an actual results orientation, with results information actually used, throughout government

3.1. Make sure that performance management is useful to programs

42. There was strong agreement at the expert meeting that performance management needs to be relevant at the agency and program level. The United States, for example, said that: "For information to be used, it must be useful." This point was underlined during the discussion at the meeting. Many other countries, including but not limited to Denmark, Korea, Hungary, Norway and Spain, also emphasised that employee commitment is essential for success. Korea, in its paper, stated that:

Performance management, including program evaluations, can have the greatest impact on final outcomes and results only when it is accepted and utilised voluntarily by line ministries themselves.

43. Performance information can be used at various levels, and for a variety of purposes. For example, one would hope that it would be useful to those directly operating programs, in order to provide guidance about what is working well and to identify areas for improvement to efficiency and effectiveness. Performance information can also be useful for program and policy development and for accountability at various other levels, including at the agency level and for the centre. In some cases, different levels of government may require differing types of information.

44. But as participants acknowledged during the ensuing discussion, if those at the program level do not view information as useful, the end result will entail going through the motions rather than actually acting in a results-oriented manner. Any resulting data is likely to be of questionable accuracy. Thus it is very much in the interests of the centre that there be commitment to a results-focused approach within all levels of government, including at the grassroots.

45. Yet engendering support among line managers and staff for performance management was identified as a problem area. Indeed, this may represent the major challenge that emerged from the expert meeting. Other participants seemed to agree with what Norway stated in its paper:

Feedback and actual institutional learning may prove to be one of the main challenges associated with enhancing a performance orientation. ... The use of performance information [is an aspect] that needs further development.

46. Sweden acknowledged that there is considerable resistance to performance management, and that managers and staff are not clear what the value is. Others agreed and expressed concern that there is substantial cynicism as well about the approach.

47. Thus there was widespread agreement about the need to create buy-in, support and use of performance management down the line, if it is to have any value. But there was a lot less clarity about how to go about doing this. Following is a brief synopsis of ideas that emerged during the discussions and that are supported by the literature.¹³

3.1.1. How to create support for a results-focused approach

3.1.1.1. Take a bottom up approach

48. Rather than imposing a system for all programs to follow, indicate that they should develop an approach suitable for their own situation and context, that can provide *them* with useful information that they can use themselves for reviewing the impact of what they are doing and identifying how this information can aid them in their own planning and practice.

3.1.1.2. Provide for flexibility

49. The above suggests that a primary focus should be on providing for a results-oriented approach that programs consider relevant for themselves. This has a challenging corollary, however, that in some cases may be contrary to current practices. For example, it recognises that one size does not fit all across an entire government. Each program will need to be permitted the flexibility that is most appropriate to its own situation and needs.

50. This view does not negate the requirements of central levels of government for performance information. It does suggest, however, that this information is likely to be most credible and accurate when programs can understand the reasons for these requests, and when they are permitted to provide the necessary information in formats that best reflect the nature of their activities.

3.1.1.3. *Provide for active participation of grassroots staff*

51. The evidence is overwhelming.¹⁴ The way to provide for buy-in is through active involvement. People are inclined to reject any approach that is imposed upon them. But if they are actively involved in its development, then it becomes their own. Ownership and commitment then follow.

3.1.1.4. *Do not expect perfection*

52. As discussed elsewhere, many of the countries at the expert meeting acknowledged that introducing a results-oriented approach has proved to be much harder than they had anticipated. It is no easier at the program or agency level. A results-oriented approach to management and budgeting requires both a different way of thinking as well as the application of techniques, such as measurement and evaluation, that few managers or staff are greatly familiar with.

53. Both the United States and Australia, the two jurisdictions with perhaps the most experience with a results oriented approach, emphasised that one should not expect perfection. For example, as the United States says in its report:

Expect failure, and do not always penalize it. When looking at failure, first understand its cause(s) before acting. In some instances, further failure may be avoided by spending more money or adjusting the management approach.

3.1.1.5. *Provide training, assistance and support*

54. Developing and using performance measures requires expertise and skills that few managers start out with. The experience of many different jurisdictions, such as in Norway, Denmark, the United States and others, is that training, guidance and the availability of technical assistance is required, over a period of time. This is needed at all levels within government, as well as with external agencies where applicable.

55. Without a variety of different forms of support and assistance, managers and staff are unlikely to be able to understand the potential value of a results-oriented approach, or be able to provide for effective implementation and use.

3.1.1.6. *Provide feedback*

56. Providing feedback to people and to programs with respect to their submissions is simple, inexpensive and highly effective. It also is all too rare.

57. Perhaps one of the major complaints of staff within government as well as within external agencies that are required to report to government, is that they never hear back on the information that they are required to submit. In many cases, they put significant effort into the preparation of their submissions, but receive no feedback on this information. They do not know if what they have done is what was expected, if it was considered useful — or indeed if anyone really looked at it other than to make a record that something was received. As Sweden acknowledged: “The agencies are uncertain about how the detailed information demanded from the ministries is used.”

58. Without at least some form of feedback, people eventually start to question if there is any value to performance measurement and why they should bother to put any effort into their data gathering and preparation of reports. This can be very demotivating and breed cynicism. It can reinforce the perception that performance measurement is just a paper exercise. In contrast, when people's efforts are recognised, and in particular when they can see how what they have done is actually used, they will begin to understand its value and be encouraged to carry on.

59. Feedback can take two different forms:

- Specific reactions to submissions.

At the very least, an acknowledgement and thank you should be sent promptly in response to submissions. The provision of actual comments, showing that someone has read and considered the information provided, is better.

- Demonstration of use.

This might take the form of a summary report, aggregations of submissions from different sources, and/or some indication showing how the information provided was put to use, for example in defending a budget submission, in the development of a new policy submission, etc.

60. If it is not possible to demonstrate how information actually has been used in some way, one should ask *why* it has been requested in the first place. Is the original information request still valid? If the information that is being received is not useful, why not, and what could make it so?

3.1.1.7. *Provide for meaningful stakeholder involvement*

61. In order to make performance management relevant, it is important to involve beneficiaries and stakeholders, *i.e.* concerned members of the public, in the process. This is essential in order to provide for credibility, as well as to improve the meaningfulness of the data that are collected, assessed and reported.

62. There are other reasons for citizen involvement as well. In particular, it represents democracy. As a recent OECD document observes: “Democratic process has increasingly embraced open government and an expectation of popular consultation before major policy changes.”¹⁵ Without at least some degree of public involvement, performance measurement risks becoming — or at least being viewed as — an internal bureaucratic exercise detached from what the citizenry views as important. Many of the presentations at the expert meeting acknowledged this and the importance of involvement.

63. For example, Korea pointed out that “performance management is a task that will be more meaningful with the participation of citizens.” The United Kingdom makes use of various consultation mechanisms in order to *involve* citizens in government. When there is no appropriate external stakeholder group, governments often will help create this. For example, the European Union has supported the development of an independent European Disability Forum, which acts as a resource and advocacy body and regularly advises the EU on policies affecting people with disabilities.

64. Yet the most common means used for obtaining public input appears to be customer satisfaction surveys. While these can be useful in combination with other means, surveys have severe limitations when used alone:

- Numerous technical and validity problems.

Scott Bayley of the Office of the Auditor-General, Western Australia, has recently indicated that most customer satisfaction surveys suffer from numerous technical problems.¹⁶ In particular, he indicates that the research demonstrates that “satisfaction measures generally lack validity.” In other words, the results of such surveys do not reflect how people really feel about the service in question, or as Senge¹⁷ has put it, it can lead to “looking good without being good.”

- Does not provide for meaningful involvement.

Asking people to respond to a set of predetermined survey questions is limited in its ability to provide for true stakeholder input. For example, it does not enable stakeholders to identify the issues and questions that *they* consider important. At best, it represents a form of consultation which is quite different from actual citizen involvement, as discussed above. This was a key theme at the 2000 conference of the European Evaluation Society in Lausanne (“Taking Evaluation to the People”). An OECD Policy Brief¹⁸ also distinguished among information, consultation and active participation, and identifies some guiding principles for engaging citizens in policy-making.

65. As the OECD report on public sector modernisation¹⁹ indicates, governments increasingly have been taking up the techniques of customer research. It is beyond the scope of this paper to consider the many alternative means for obtaining public sector input and providing for meaningful involvement.

3.2. *Creating a results-focused culture*

66. A strong consensus appeared to emerge from the discussions at the expert meeting about the importance of a results-oriented culture. This was stated in many different ways, and may represent the most important theme arising from the meeting. When there is an actual culture that values an orientation on results, people take a results-focused approach not because it is mandated, but because they want to, because they see the value in this themselves, in many cases because they cannot conceive of doing things any differently. The *thinking* process itself is critical. The value and beliefs inherent in a results orientation are internalised and acted upon as a matter of course.

67. There appeared to be general agreement that performance management can only be effective if people working within the system actually *believe* in the value of a results orientation. With a true results focus, a results focus permeates everything that one does, for example, with managers and staff constantly asking themselves what the benefits of their activities are supposed to be, how they know, and what they can do to be more effective. Otherwise, what takes place too likely would be mainly a paper exercise, with people going through the motions rather than really acting in a results-oriented manner.

68. Performance management most frequently is considered part of the formal, structured organisational system. Yet virtually all studies of how organisations *really* work emphasise the importance of the informal organisation. Much of the real communication within organisations happen informally,

rather than along the lines indicated in organigrams. The informal organisation is where many things really get done. It also can be the major source of motivation for professionals, and frequently for rewards as well (e.g. respect from one's peers within or outside government, which can be more important to knowledge workers than how their own hierarchy views their work). The only way a results orientation can permeate the informal organisation is through culture. Otherwise, performance management risks representing a model separate from how things really get done.

69. As the United States indicated: "Do not underestimate the sense of individual pride and satisfaction that government managers and workers gain from doing their jobs well." Most public servants are very committed to their work and they *want* to make an impact. That is why they are in the public service in the first place. If they can see how a results-oriented thinking can help them in this regard, rather than serve just as an external control mechanism, they are much more likely to buy into the concept.

70. How can a results-focused culture be created? No simple answers to this emerged from the expert meeting, despite a recognition of its importance. There was widespread recognition that culture, by its very nature, cannot be changed overnight. One must be realistic, recognising that it requires a long-term approach. But it is possible to monitor progress, to identify intermediate outcomes, to stimulate activity where interest is greatest, and to reward those who demonstrate thinking and action in results terms.

71. Neither can a results-focused culture be mandated, given that its essence is how people think and their values. Three related strategies emerged from the discussion, which are briefly discussed below:

- a human resources management approach;
- leadership;
- priority to learning.

3.2.1. *Human Resources Management*

72. It is all too easy to confuse performance *management* with performance *measurement*. They are not the same, even if they frequently do share the same acronym (PM in English). In fact, measurement is just one of many *management* tools. It should not be confused as an end in itself. Similarly, performance (or results-based) management frequently is viewed mainly as a matter of formal data collection and reporting systems and formats.

73. But such a view ultimately would be self defeating. As we have seen above, actual implementation, rather than the mere illusion of a true results-oriented approach is dependent upon the thinking process, along with actions, of people. A background paper presented by the Netherlands reinforced this point. It emphasised that the stimulation of *people* within the organisation is most important, and that management involves the work of men and women. France, in its presentation, emphasised the need to focus on human resources management. Other presentations also spoke to this point. The OECD paper prepared by Pollitt²⁰ on integrating financial and performance management also indicated that human resources management is of "critical importance to a well functioning resource management system."

3.2.2. *Leadership*

74. One can order people to undertake specific activities. But it is impossible to order or to direct people how to think or what to believe. Indeed, this is most likely to be counterproductive. As Korea for example indicated, one must take extreme caution not to impose: a performance-oriented approach must be accepted and used voluntarily. Thus a top-down command-and-control approach to management may be effective in some instances in inducing short-term behaviour change, or perhaps for production type activities. But is not suitable for creating a change organisational culture, or for getting the best from knowledge workers.

75. What is required to create culture change, and to motivate knowledge workers in particular, is *leadership*. As Denmark indicated: “Leadership is regarded as an essential part of communicating the vision, mission and value statement.” Effective leaders work not by ordering people what to do, but by instilling a desire in them. Leaders work at the basic human level by guiding, by providing stimulation and motivation, by transmitting a vision that people *want* to buy into. Effective leaders work with their staff to develop a common understanding of expectations and then clear hurdles to enable them people to take effective action. Effective leaders provide resources, reinforcement, encouragement and support. They recognise that most people want to do good, and reserve sanctions only for exceptional situations. A results-focused approach cannot be mandated — but conditions can be set in place where this can germinate and develop.

76. To be sure, there are debates about how best to provide leadership and the extent to which leaders are born *vs* trained. Yet the above represents mainstream management and human resources thinking. A number of the participants at the expert meeting seemed to recognise this. Nevertheless, most of the formal papers still seemed to take a control rather than a leadership model for the implementation of a results-focused approach. There are, for example, many references to terms such as “control”, “compliance” and “discipline”, and in contrast, very few to “leadership”, “inspiration” or “motivation”.²¹

77. There is no question that major responsibility for the overall vision, priorities and direction comes from the top in government, from elected representatives and their designates, just as it comes from the boards of directors and chief executive officers in the private sector. But how can this vision be instilled throughout the organisation and transformed into the necessary actions to bring it about? As indicated above, the evidence is quite clear about the limitations of command-and-control, in particular among knowledge workers. If a results-focused culture is to be created, and if there is to be buy-in to a results-focused approach, there needs to be more recognition of the paradox that one only can obtain more control through less control.

3.2.3. *Priority to learning*

78. There appears to be at least some interest in using performance management as a tool for learning. Spain, for example, emphasised the importance of creating a performance culture, and how an increasing capacity for learning is needed in order to bring this about. Norway indicated that: “Feedback and actual institutional learning may prove to be one of the main challenges associated with enhancing a performance orientation.”

79. What does it mean when results information is used for learning? In this situation, managers and their staff view results information as useful feedback about what is and is not working well. Monitoring and evaluation are used to probe into potential problem areas, to understand the reasons why some things may not be so effective, and what forms of corrective action may be appropriate. Thus a focus on learning is consistent with a focus on program improvement, and on identifying new approaches that may be needed to addressing recurring or new needs.

80. With a learning focus, managers and staff constantly seek out information about what they know is not working, looking for ways to improve. As Australia, for example, indicated, mistakes should not necessarily result in punishment, as they can represent learning opportunities. This is consistent with the literature on how to promote innovation.²²

81. But placing too much attention on whether or not pre-determined targets have been achieved or not can be inconsistent with learning, especially if sanctions (*i.e.* blame) is attached, or even anticipated. Such a “blame” approach is counter productive. Instead of encouraging managers to ask the difficult questions, it punishes them for doing so. It almost forces programs and managers to justify what was done rather than to acknowledge what is not working well and areas for improvement.

82. Accountability, in the context of performance management, is most typically viewed as accountability for meeting pre-determined targets. Instead, would it not be more appropriate to hold programs accountable for demonstrating a results-focused approach? Under this definition of accountability, programs would be accountable for asking the difficult questions, for doing evaluation, for using and applying the findings, for demonstrating learning and for acting upon this.

83. This view of accountability is consistent with the OECD definition of accountability, and with how it has been proposed within Canada:

*To encourage and support managing for results, we need a new view of accountability that acknowledges this more complex management world. ... Accountability for results or outcomes asks if you have done everything possible with your authorities and resources to effect the achievement of the intended results, and if you have learned from past experience what works and does not work. Accounting for results of this kind means demonstrating that you have made a difference; that through your actions and efforts you have contributed to the results achieved.*²³

4. Data quality: Key issues and potential solutions

84. Use of performance measures or indicators is a basic component of results-based management. It is predicated on the assumption that programs can and do produce meaningful, accurate, reliable and valid data about performance. Yet there is extensive evidence indicating a real danger of misrepresentation and misuse of performance measures.²⁴ This evidence suggests that *misuse may be commonplace rather than the odd exception*. It is a major reason for the demise of previous approaches to performance management over the decades, in different parts of the world, and in both the private and public sectors. Without attending to the limitations of performance indicators and problems associated with misleading data, results-focused management risks irrelevancy, or worse.

85. Many of these data quality problems potentially can be avoided, or at least taken into account in the interpretation and use of the data. But in order to be able to anticipate these potential problems and to be able to address them or take them into account, one first must acknowledge explicitly that these problems may exist. Yet there was hardly any mention of this potential in the country papers or presentations. Almost all papers seemed to assume that reported data can be taken at face value, without question.

86. The appropriate use of performance measures is important for those at the agency level charged with implementing and using a results-focused approach. But it clearly is also a concern for those in central agencies, such as Ministries of Finance or equivalent, with strategic responsibility for management and budgeting across government.

- If the data that are reported do not reflect actual performance, they are meaningless or worse. It would be misleading or even dangerous to use them for policy- or decision-making.
- In mandating or promoting a performance management approach across government, it is important to ensure that the approach indeed does what it says it does. It is of particular importance to avoid perverse effects, where such an approach may result in the opposite of what was intended.
- At the very least, being aware of potential problems and pitfalls can suggest potential questions and issues to monitor on an ongoing basis and to consider when reviewing reports from agencies.
- Similarly, awareness of potential difficulties, and how these can be addressed, is a basis for providing guidance to agencies about effective implementation and use.

87. There is an increasing literature²⁵ documenting limitations and problems associated with performance measurement. The following brief discussion highlights just some of these issues that have been identified, and suggests how these can be avoided or addressed.

4.1. *Some key issues with performance data*

The government is very keen on amassing statistics. They collect them, add them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, who just puts down what he pleases. — attributed to Sir Josiah Stamp.²⁶

The table at the right summarises many of the key issues that have been raised in the literature about performance measurement data. Four of the key considerations are discussed below.

4.1.1. *Misrepresentation*

88. Very recently, an article in *Nature* indicated that statistics reported by China regarding the fish catches by its fishery had been grossly distorted. “Apparently, under the Communist system of matching results with plan, the same bureaucrats were responsible for not only counting the catch but also meeting targets to increase it — so they simply exaggerated the count to match their allotted goals.”²⁷

89. These are the same data used by the United Nations to assess global fish catches and to estimate overall fish populations in the seas. Because China forms such a large part of the total, it now appears that the world’s catch has actually declined by an average of 360 000 tons per year, instead of rising by 330 000 as the erroneous data had suggested. These spurious data have led to inappropriate conservation policies, overfishing, and unwise investments in fishing fleets.

90. It is becoming increasingly apparent that distortion of performance data in order to meet targets is endemic in the private sector as well. For example, a recent article in the *Harvard Business Review* starts out as follows:

*Corporate budgeting is a joke, and everyone knows it. It encourages managers to lie and cheat, lowballing targets and inflating results, and it penalizes them for telling the truth. It turns business decisions into elaborate exercises in gaming. It sets colleague against colleague, creating distrust and ill will. And it distorts incentives, motivating people to act in ways that run counter to the best interests of their companies. It’s the use of budget targets to determine compensation. ... Only by severing the link between budgets and bonuses — by rewarding people purely for their accomplishments, not for their ability to hit targets — will we take away the incentive to cheat.*²⁸

Dangers of misuse of performance indicators

- goal displacement;
- varying interpretations of definitions;
- distorted or inaccurate data;
- meaningless and irrelevant data;
- cost shifting to other areas vs. actual cost savings;
- critical subgroup differences are hidden;
- do not take into account the larger context;
- same limitations as other objective-based approaches to evaluation;
- useless, by themselves, for making decisions (e.g. do not provide information about the reasons for given findings);
- can result in less, rather than more, of a focus on outcome.

91. Performance data in the non-profit sector also can be distorted. For example, the newsletter of an adult literacy (basic education) association acknowledged that: “Adult literacy workers are so overworked, underfunded and overregulated that many burn out and quit, while others ‘fudge’ their performance figures.”

92. One can manipulate numbers in a way that while not necessarily dishonest, can be equally misleading. For example, *The Economist* has indicated perverse effects of “stretch goals”:

*In good times these ambitious targets encourage managers to fiddle the figures so that they are achieved, just. In an economic downturn, when the targets move completely out of reach, those same managers are persuaded to fudge the figures in the opposite direction. They bundle as much bad news into the current period as possible, aiming to make their targets easier to achieve in the next.*²⁹

93. Performance data do not have to be wilfully distorted to misrepresent reality. A commonplace situation is when staff supplying data pay little attention to what they record. An example of this might be busy social workers, who focus mainly on their work with clients and place low priority on what they view as burdensome and irrelevant “paperwork”. Nevertheless, in terms of the resulting data quality and what it represents, the end result can be the same.

4.1.2. *Meaningless or misleading indicators*

94. Performance data can be meaningless or misleading for a variety of different reasons. Most frequently, this is not intentional on the part of those who supply the data. Yet it makes the data no less useful than in the above situations where there may be exaggeration or even outright deceit.

95. There can be a variety of reasons why reported data can be meaningless, or worse. Consider the case of a politically sensitive program in Canada, where courthouses were required to submit weekly performance data. Yet when the clerks who prepared these reports were interviewed, they explained how it was impossible for them to provide information for any of the required categories. When they were asked how they completed and submitted their weekly statistics, they responded: “We put down something that sounds reasonable.”

96. Frequently, there are varying and conflicting definitions, even of some seemingly commonplace items as “numbers of “clients”. When such data are aggregated, they have no meaning. For example, following changes in response to criticisms of tardy deliveries, a military procurement was able to report 99% on-time delivery. But while this technically may have been accurate, it subsequently transpired that each item shipped was recorded equally, no matter how large or small. Thus the 99% figure could be achieved by shipping a box of 100 bolts on time, even if delivery of a crucial jet engine was considerably delayed.

97. Or consider another example: health vs. reported health. The availability of health services paradoxically can result in an increase in reported health problems, making it appear that health can be worse than in areas with few health services. The creation of child abuse prevention programs invariably results in an initial *increase* in recorded incidents of abuse. For similar and other reasons, reported crime statistics may bear little relationship to the extent of actual crime.

98. Data such as the above may be meaningless. But because they can be made precise, however spurious, they can *appear* meaningful and take on a life of their own.

4.1.3. *Goal displacement*

99. Goal displacement occurs when indicators become the objective, where the focus is on “meeting the numbers” rather than doing what the program was created to do or improving actual outcomes. Thus it represents the distortion of program activities. Examples of goal displacement are legend.

100. For example, it recently came to light that waiting lists for surgery in Britain’s National Health Service were reduced by treating those easiest to service, *e.g.* with less serious conditions, even if it meant that people whose conditions were life threatening continued to wait for treatment. Perhaps the classic example of goal displacement and the distortions that can result is the infamous Vietnam body counts, where counts of “enemy” dead were used by the United States as an indicator of progress in fighting the war. Goal displacement also frequently comes from “creaming”. This involves serving the easy to serve rather than those most in need, such as selective entry of clients into a job training or placement program.

101. Goal displacement may, but most often does not, involve deceitful behaviour. How can one fault a manager for giving the most attention to the priorities that have been identified and agreed upon? As *The Economist* had indicated: “If public servants are asked to focus on one measure, they will (rightly) ignore the others.”

102. Performance audits may be able to identify at least some cases where numbers have been fabricated. Possibly in some cases they may be able to identify where the numbers provided have no basis in fact. But they usually are not in a position to be able to identify situations where public servants are doing their jobs, working to meet the targets that have been set out to them, even if this may be inappropriate or have negative consequences.

4.1.4. *Inappropriateness of reductionist approaches*

103. Many government interventions typically are complex in nature, involving many interacting variables, with progress likely to be irregular over a period of time rather than regular and predictable in nature. It does not always make sense to attempt to represent a complex initiative by just a very small number of quantitative indicators. There typically is an inverse relationship between what is important and what is easiest to count and to measure.

104. As Handy says, traditional measurement systems “conditioned by the philosophy of the audit ... to count only what they can put their finger on and cross their hearts about” results in more attention to efficiency than effectiveness and to lack of attention to what is really important, even if difficult to quantify. Mintzberg³⁰ has noted that: “Assessment of many of the most common activities in government requires soft judgment — something that hard measurement cannot provide. ... Measurement often misses the point, sometimes causing awful distortions.”

105. In an article entitled “The Earnings Game”, Collingwood³¹, senior editor of the *Harvard Business Review*, indicates how too much attention to quarterly earnings in the corporate sector distorts corporate activities, with more attention placed to meeting earnings expectations than on return to shareholders or profits. Collingwood indicated that: “The central problem of such estimates ... is that they attempt to reduce a complex enterprise to a single number of minimal predictive value.”

4.2. *How to provide for appropriate use of performance measures*

106. As indicated at the beginning of Section 4, problems with the quality and meaningfulness of performance measurement, such as those indicated above, do not appear to be the odd exception. They are commonplace. But there are many strategies that one can take to minimise these difficulties and to avoid misuse of performance data. The box on the right lists a number of strategies. Some approaches are briefly discussed below.

107. Perhaps the most important thing is to bear these factors in mind, to anticipate potential dangers and identify in advance how they will be addressed. Similarly, when reviewing findings, recognise that data do not speak for themselves. As Section 2.3 suggested, one should consider alternative plausible explanations when interpreting findings and seek answers to these as need be.

4.2.1. *Do not tie decisions on compensation or budgets to achievement of targets*

108. Linking compensation of managers, or of future program budgets, to the achievement of pre-set targets can make goal displacement practically inevitable. This places undue pressure on a manager to “meet the numbers”, even at the exclusion of other factors that may be more important. Ironically, while the private sector frequently is cited as a model for linking pay to performance, there is increasing recognition of the perverse effects that this can cause.³² For example, Kenneth Lay, former chief executive of Enron, reportedly received US\$123M for achieving his target, at the same time that the company was virtually defunct.

In order to use indicators appropriately, one should:

- Distinguish between performance *data* and performance *information* — no single or limited number of indicators can ever give the full picture.
- Identify limitations and provide qualifications.
- Use indicators as part of a broad evaluation strategy, complementing with other information, including qualitative information such as “stories”.
- Use indicators to raise questions for exploration rather than to provide the definitive “answer”.
- When reporting, provide more rather than less information than required, to help put indicator data into context.
- Constantly reassess validity and meaningfulness.

The process of developing indicators should include:

- The involvement of stakeholders in the development, interpretation and revision of indicators.
- Allocation of time and resources to the development of indicators.
- Provision of training and expertise.
- Thinking about potential forms of misuse in advance.
- Pretesting, testing, review and revision of indicators on an ongoing basis.

109. OECD countries are divided in their approach to “performance-based pay”. A small number of countries practice this approach with at least some of their managers, while a few others are considering this. When it does occur, it rarely is at the magnitude as used within the private sector. One would like to think that people working within the public sector are motivated by more than just remuneration. Nevertheless, as for example one country report explicitly states, the objective of this approach is “to create an incentive to achieve set targets.”

110. Probably most OECD countries represented at the expert meeting appear to recognise the dangers and limitations of tying performance explicitly to the realisation of targets. Some have pointed to other factors that must be taken into consideration. For example, the United States indicates in its report that one must understand the causes of poor performance. For example, in some cases it might be a result of underfunding. Thus it is not appropriate to base compensation just on meeting targets.

111. This does not mean, of course, that performance should not be taken into account in deciding upon managerial effectiveness, promotions and remuneration. The issue is tying compensation specifically to target achievement. Boyle³³ discusses a variety of alternative models in the public sector of rewarding performance.

112. One should bear in mind that manager and staff act in accordance with their expectations, even if this may differ from reality. Thus if they feel that their program budgets can be threatened if they do not achieve their targets, even if this in fact is unlikely to happen, they may be tempted to “adjust” their figures accordingly. This pressure can be especially strong for external agencies funded by contract.

4.2.2. *Use quantitative measures only when they are appropriate and meaningful*

113. Quantitative performance measures are appropriate in some circumstances. But as has been indicated, they are not appropriate in all situations. The literature suggests that they tend to be most appropriate for routine, predictable operations. They are unlikely to be applicable for more complex, innovative or ever-changing initiatives or activities. One should not mandate the use of performance measures in all circumstances.

114. For example, as both Mintzberg³⁴ and Pollitt³⁵ in his OECD report indicate, performance measures may be appropriate for tangible, standardised activities, such as issuing licences or passports. They are not so likely to be appropriate in more complex situations, such as for tangible individual services (e.g. health care) or with respect to non-tangible ideal services, (e.g. coordination of other activities or provision of policy advice). The Netherlands report indicates that:

Quantitative and annual measurements are easier to attain for processes with a more implementing character. The more it concerns policy matters, the more it shifts towards multi-year quantitative and qualitative evaluation research.

115. Hungary indicated that: “Performance-focused management is applicable mainly in fields where performance can be measured in monetary terms.”

116. It is interesting to note that the case studies that were provided as part of some of the country reports are generally consistent with the above. Most are of routine operations, such as tax collection or customs. Of the two exceptions, one is of the Sure Start Program in the United Kingdom, where it is premature to make any judgement and where there are decidedly mixed views of the approach that is being taken. It is also significant that the program is also subject to a comprehensive program evaluation. Thus the performance measures form only one component of the overall assessment of performance.

117. The other exception is in Australia. But again, the report strongly suggests that performance indicators in the social policy area can only give part of the story. The report suggests that: “Some caution needs to be exercised in the use of performance targets in the social policy field,” and that: “In some respects the term ‘target’ is less relevant in social policy.” Again, there appear to be differing views about the applicability of targets, with those in the implementing agency being most sceptical. And evaluation is also used along with performance targeting.

4.2.3. *Make performance measures useful to programs in their own work*

118. As Section 3.1 indicated, perhaps the most important way to get meaningful performance information is for it to be meaningful and relevant to those who must produce the information. If staff must collect information just for reporting “up the system” that has no meaning to them, they are unlikely to give this a high priority or do little more than just go through the motions. Data accuracy is likely to be questionable at best in such circumstances.

119. On the other hand, if managers and staff can see how performance information can help *them* in their own planning and practice, they are likely to put more effort into ensuring its relevance and validity. Thus the best way to provide for meaningful results information at any level within government is through the creation of a results culture, and for providing assistance to enable managers to develop and to use appropriate monitoring and evaluation techniques that they feel would help them in being as effective as possible.

4.2.4. *Use performance measures to raise questions rather than to provide definitive answers*

120. As has been indicated, performance measures are not sufficient, in and of themselves, for answering questions about the overall performance of complex questions or programs, or for deciding upon managerial effectiveness. They *can*, however, play an important role in *raising* questions and in focusing attention on issues that require exploration through other informal and/or more formal means. In particular, as Section 6.1 discusses, performance measures can serve as a means of identifying where more comprehensive evaluation approaches would be most useful.

4.2.5. *Recognise that it is not easy*

121. Many of the countries, in their papers and presentations, such as Sweden and Spain, acknowledge that introducing performance measurement proved to be much more difficult than they had anticipated. This in turn has a number of implications.

- Significant *assistance and support* are needed.

As Section 3.1 has indicated, the development and use of performance measures requires skills that few managers start out with. Without significant support and assistance of various forms, the quality and accuracy of performance measures, and their utility, is almost certain to suffer.

- Take a *dynamic* rather than a static approach to performance measurement.

Both supporters and detractors of performance measurement acknowledge that measures are never right the first time. Australia, for example, indicates that it takes two to three cycles to develop a meaningful system. The United States advocated proceeding slowly. The United States and Denmark are two examples of countries that, at least to some extent, started by pilot testing their approaches to performance management.

- Always pre-test.

Experts always pre-test potential measures, often multiple times, recognising that they are rarely perfect to start with. It is the neophytes who sometimes think that this is not necessary. Without cycles of pre-testing and revision, measures are almost certain to be flawed.

- Recognise that indicators have a limited half life.

Indicators can become out of date very quickly for a variety of reasons. The literature suggests as well that indicators that remain unchanged are most susceptible to corruption.

- Expect measures and targets to change.

If programs are responsive to changes in needs, environmental conditions and circumstances, then program approaches along with performance measures *should* change from time to time. One might well be suspicious if program indicators and targets remain unchanged for any length of time.

5. The role of central leadership and support

5.1. Top-level support

122. As Norway indicated in its presentation, commitment and involvement from top management within government as well at the agency level is a pre-requisite to the success of a results-focused approach. Denmark also highlighted the importance of leadership as essential in getting across the vision. Spain, pointed out that political level support for reform is also essential.

123. In this respect, developing support for a results-oriented approach to management and budgeting is no different than with respect to any other major government initiative. There always seem to be a multitude of competing directives and priorities. If top management does not indicate that a results-focused approach is of high priority to it, then why should anyone else invest in the approach? Support can come through communications. But action is even more important. When people see that top management really

is taking a results-oriented focus in what it is doing itself, then the importance of this approach is most likely to permeate throughout the organisation. Conversely, ignoring results information will give a very different message about how important it really is.

5.2. *A central unit with lead responsibility for promoting a results-oriented approach*

124. All the representatives at the expert meeting seemed to agree that while it is appropriate for the centre to mandate a results-oriented approach, actual implementation needs to be at the agency level. This is necessary in order to develop an actual approach that would be most suitable for each particular situation. Nevertheless, a central unit, in a prominent place in government, is essential in order to set expectations and to drive and support a results-oriented approach across government. Such a unit is also essential in order to establish overall policy, to review and to adjust the approach as needed, to identify how results-oriented information can be used in government-wide policy and decision-making, and to provide direction, assistance and support as needed in order to facilitate the actual development and use of a results-oriented approach.

125. It is important to recognise that performance information can be qualitatively different in nature from financial information. For example, financial information is always quantitative, reducible to a single unit of measure (*i.e.* the monetary currency), specific and objective in nature (although it is increasingly apparent that even with budgetary information, there can be many subjective aspects, such as about what should be included and in what ways). Performance information, in contrast, rarely can be as specific or exact. At best, performance indicators can give a “window” on the underlying reality rather than represent it fully. As discussed earlier, the higher up the results chain one goes towards the most important questions about outcome, the more difficult it is to specify things with complete certainty.

126. Furthermore, performance information is most useful, and most likely to be accurate, when it does not come across as adversarial, and when it can be used for learning and to provide guidance for future direction. To reinforce this message, it would be preferable to separate lead responsibility for promoting a results-oriented culture in some way from budget control functions. Responsibility for a results-focused approach in most OECD countries appears to reside within the Ministry of Finance or equivalent. This can be appropriate, given the need to emphasise the importance of a results-focused approach, as long as the differences between financial and performance information are recognised.

127. It is also important that a unit promoting a results-oriented approach be staffed appropriately. The backgrounds and skills required of managers and staff normally should be in disciplines such as evaluation and human resources, rather than in accounting or economics, as would be more appropriate in many other areas within a finance department. As well, it would be advisable to have at least some people with recognised line experience about how things *really* work in practice at the grassroots level. In some cases, this can be done by secondment or rotation.

128. It would also be useful to have a government-wide advisory body of some form, with grassroots representation from across government. This can help with communications and to enhance the credibility of the unit. It can also help ensure that the approaches it mandates or advises are realistic and relevant, and that it takes into account the concerns and challenges faced by those in the field who need to put a results-oriented approach into effect. It also may be able to help identify potential problems and limitations with data quality, so that performance information is more likely to be used appropriately.

5.3. *Capacity building in performance management*

129. As indicated earlier, a key learning from the expert meeting is that implementation of a results-oriented approach is not easy and has proved to be more difficult than expected. It will not happen by itself. Expertise and a body of knowledge are essential. Without these, the dangers of misuse of performance data are also heightened.

130. Most OECD countries acknowledge that there is limited expertise in performance management and evaluation, at central levels as well as within agencies. Staff in central agencies need a better understanding about the nature of performance, the strengths and limitations of performance measurements, the potential of evaluation, and how to use performance information most appropriately. Staff in agencies require expertise in how to plan and implement appropriate approaches to performance monitoring and evaluation, and to apply the findings. Training and support of various forms are needed. Some countries are taking action along these lines. *Without the creation of capacity in performance management, a results-oriented approach likely is doomed to failure.*

131. It was also recognised at the expert meeting that culture change is a long-term endeavour. The experience across many different jurisdictions is that long-entrenched thinking largely in terms of inputs and activities cannot change to an outcome focus easily or quickly. To address this challenge, as well as to help build a culture of evaluation within government and to provide for long-term sustainability, Malaysia has taken a long-term approach to capacity development. For example, it has made available very extensive training, including a certificate program in evaluation from RMIT University in Australia, to a limited number of staff in each government agency. These people are then able to act as facilitators to promote and nurture evaluation and the development of a performance culture within their own organisation. Phase 1 involved training of some 15-30 facilitators within 11 agencies. The second phase is now under way, involving another 10 agencies.

5.4. *Communications*

132. Effective communications within government, and with any relevant external agencies, is critical to effective implementation and use of a results-focused approach. Many governments are providing extensive information about performance management systems and how they are expected to be used.

133. It is particularly important to check, rather than to assume, that the message that is being sent out is the same as what people are hearing. For maximum effectiveness, two-way communications is needed. People act according to their perceptions, which is not necessarily the same as what is really the case. Particularly with respect to performance information, there can be fears that this information may be used against managers and programs, for example in cutting budgets or even terminating programs if performance targets are not achieved, whether these fears are realistic or not. Such a situation can heighten inappropriate use of performance management, such as the dangers of goal displacement and data distortion that have been discussed in Section 4.1.

134. Two-way communication is also important to make sure that central units are aware of what is really happening with performance management at the grassroots. Given the importance of the informal organisation, it may be necessary at times to go beyond formal hierarchical reports to get the full picture. There can be a lack of credibility which can inhibit results-oriented actions when central issue proclamations about how things work, when people down the line feel that the reality actually is quite different.

135. In fact, communications need not be just between the centre and the departments or agencies. Norway, for example, has facilitated the establishment of networks and forums to provide for the sharing of experiences to facilitate inter-organisational learning and discussion among practitioners.

5.5. “Take your own medicine”

136. Countries at the expert meeting were challenged by other participants to “take your own medicine” — to evaluate the impact of their own results-based initiative, just as they are urging agencies to do so with respect to their own policies and programs. Given the extent of attempts at reform in so many different countries over the decades, there has been surprisingly little evaluation of these initiatives.³⁶ Previous OECD reports have confirmed the lack of evaluation of reform activities.³⁷ e.g. one recent document, reviewing public sector modernisation over the last 10 years, identified the need for “better evaluations of new practices”, along with the “the need for more empirical information about whether innovations have worked or not.” This paper adds that many important public management changes that have been presented as best practices later appeared to develop undesired and unanticipated consequences.

137. Many of the country papers and presentations at the meeting confirmed there is very little actual evidence about the effectiveness of results-focused approach. Section 2.2 provided citations from some country reports acknowledging that reform activities, in spite of their stated intent, have not focused on outcomes, Spain is perhaps more direct than others in acknowledging that:

For almost two decades, we have seen many innovations within management, budgeting and auditing without there having been attempts to assess the real effects of such reforms. Their limited impact is a widely accepted fact.

138. Proper evaluation of government-wide results-based initiatives can have a number of advantages:

- It represents leadership by example, modelling what one preaches to government departments and agencies. In this respect, it can help stimulate and reinforce implementation of a results-focused approach throughout government.
- It has the potential to provide documentation of the benefits of such an approach. This information in turn can be used to generate additional support, at the senior management and the political levels, as well as down the ranks.
- A formative or learning approach to evaluation can identify what is working well and what is not and why. This can provide invaluable information about the strengths and limitations of the approach that has been used and how it can be improved. Given the number of challenges to results-focused management that have been raised by all the participants at the expert meeting, such an approach to evaluation would appear to be a priority.

139. *How* evaluation is done can have a major impact on how it is received and viewed. A participative approach, involving stakeholders from inside and outside government, can help both with credibility and meaningfulness, to ensure that the right questions are considered. Similarly, actual conduct of the evaluation should be independent of the central unit responsible for the results-focused approach.

5.6. *Audit*

140. Traditionally, the role of auditors has been to attest to the integrity of the financial accounts and expenditures. This is still the primary focus of auditing. But increasingly, auditors are engaging in performance or “value-for-money” audits. These activities can go well beyond the examination of financial data.

141. Increasingly, supreme audit institutions are taking an active role in providing quality assurance for performance reports, and perhaps for the integrity of the overall approach to performance management and the resulting performance data. For example, Denmark notes in its report that its National Audit Office audits the annual performance reports submitted by agencies, and assesses the reliability and documentation of the information provided in the annual reports. Section 4 indicated that, at least to some extent, performance audits may be able to identify instances of the types of data quality problems that have been identified.

142. Thus there clearly is a role for audit to play with respect to performance management. As performance information takes a higher profile, an external seal of approval attesting to the quality of this information would seem to be a useful function and an appropriate role for supreme audit institutions. This, however, can raise a number of issues, such as the following.

143. **Legal status.** For example, what is the legal status of performance information, when this is reported to Parliaments along with budget information? All OECD countries make a distinction between material included in the budget for information and material included as part of the legally binding vote. But the closer that financial and results information and targets are aligned, the more the latter can be regarded as part of the appropriation. This in turn can lead to increased pressure for programs to meet targets, with all the perverse effects that have already been highlighted.

144. **Auditor expertise.** As noted earlier, performance data are quite different in kind from the financial data that auditors traditionally have been trained to deal with. This in turn has implications for the selection and training of audit officers with appropriate backgrounds in evaluation and the social sciences rather than in accounting. It should be noted that this is indeed the case in the audit offices in many jurisdictions.

145. **Appropriate scope for audits.** What level of analysis is it appropriate for auditors to take in performance audits? For example, should audits focus primarily on affirming the use of appropriate management approaches, the consistency of the procedures used to gather performance information with those used elsewhere, and the consistency of what is included in reports with the available data? For example, on page 16, this report suggested that programs should be held accountable not for target realisation, but instead for asking the difficult questions, for doing evaluation, and using and applying the findings, for demonstrating learning and for acting upon these. It might be an appropriate role for performance auditors to ascertain if this has been done.

146. But in some jurisdictions, there is the potential for auditors to go further, for example in making judgements on the quality and appropriateness of the targets and the results obtained. Is this an appropriate role for audit? Questions have been raised about the appropriateness of auditors delving into such detail, given that they are not usually equipped with content level expertise and awareness of the context in order to be able to make appropriate judgements such as these.

147. Dangers of an excessive judgemental focus. Audit reports most typically are judgemental in nature. As Boyle, for example, has observed: “Performance auditors may have a tendency to focus on weaknesses and have been accused of producing non-balanced reports.”³⁸ As we have seen, a blame approach based upon the principles of compliance and blame, such as have been associated with traditional audit, can be counter productive to the development of a results-focused culture. Power³⁹ has shown how the growth of what had been termed the audit society, with too much emphasis on checking and compliance with rules, can interfere with an emphasis on effectiveness and on innovation.

148. **A supportive role for audit bodies?** In some situations, auditors attempt to play an educational role, as a complement to their main stock in trade of conducting audits and providing judgements. The National Audit Office in the United Kingdom, and the General Accounting Office in the United States, are but two examples of audit bodies that make available supportive guides on a variety of topics including performance measurement, survey design and implementation, etc. These guides can represent useful educational resources. But this approach also is not without its critics. For example, it is appropriate for the same body responsible for making judgements also to provide guides such as these? In many cases, these guides have been prepared in areas where there are no clear answers and where there are other resources that might advocate somewhat different approaches. Are programs subject to a negative audit if they do not follow the advice of the guides issues by their own audit offices?

149. In short, it is clear that audit can have a significant role to play with respect to performance management. But it also is not without its ambiguities and dangers. Depending upon how various competing factors are balanced, audit can support — or inhibit — implementation of a results-focused approach.

6. Strategic considerations

150. Participants at the expert meeting acknowledged the complexity of many of the issues and the dangers of a piecemeal approach to performance management and budgeting. Many countries are also grappling with questions about how performance information, if at all, can be used with respect to external contracting for services, the budgeting process and reporting to Parliaments. Issues such as these are considered below.

6.1. *Flexible and integrated approach to monitoring and evaluation*

151. A significant theme arising collectively from the papers and presentations at the expert meeting is a growing recognition that no single approach to results-focused management and budgeting is appropriate in all circumstances. This can be true at both the macro- and the micro-levels. The initial experiences of a number of countries has been that whatever approach is taken must be flexible and adaptable. Denmark, for example, has indicated that through the pilot projects of its performance management approach, it learned that the approach must be adapted to fit each ministry.

152. Similarly, rather than try to force programs into a pre-established framework and method for assessing performance, it is essential to adjust the method to suit the circumstance rather than vice versa. Otherwise, one risks coming up with irrelevant information, as well as increasing cynicism about the entire process. As the European Union has said:⁴⁰ the Golden Rule with respect to methods and techniques is that

“there are *no* golden rules.” For example, as has been discussed earlier, quantification is appropriate in some circumstances, but not in others. Generally a balanced approach with a mix of different types of information is needed. As Section 2.3 has indicated, one can use performance measures most sensibly by seeking out whatever form of additional evidence is most appropriate in order to address alternative explanations arising from the initial data.

153. In particular, there appears to be increasing awareness that as Section 3.2 has indicated, performance monitoring and evaluation can provide different types of complementary information. While the major emphasis among the countries making presentations is on monitoring data, many also make use of evaluation, at least to some extent. An effective performance management system, in particular one that can provide a focus on outcomes, requires both.

154. Nevertheless, there are varying degrees of understanding among public officials and staff about the differences between monitoring and evaluation. For example, some people think that tracking indicators *is* evaluation.

155. UNICEF⁴¹ defines *monitoring* as:

The periodic oversight of the implementation of an activity which seeks to establish the extent to which input deliveries, work schedules, other required actions and targeted outputs are proceeding according to plan, so that timely action can be taken to correct deficiencies detected. ‘Monitoring’ is also useful for the systematic checking on a condition or set of conditions, such as following the situation of women and children.

156. It defines *evaluation* as:

A process which attempts to determine as systematically and objectively as possible the relevance, effectiveness, efficiency and impact of activities in the light of specified objectives. It is a learning and action-oriented management tool and organizational process for improving both current activities and future planning, programming and decision-making.

157. In essence, monitoring, such as the use of performance measures or indicators, involves assessing performance against pre-determined targets, usually using readily available data. Such data can be invaluable at indicating the extent to which things are going according to plan. Data obtained through monitoring, however, rarely is sufficient to get at outcomes, and used alone, is not capable of identifying attribution or causality (*e.g.* to what extent can one say that a program activity is directly responsible or not for the given results?). Monitoring information often can best be used to direct management attention and to raise questions that should be explored in greater depth through evaluation.

158. Evaluation has the potential to identify what *actually* has happened as a result of a program or initiative, whether planned for or not, including unintended outcomes and effects that often may be more significant than the stated objectives. In particular, evaluation has the potential to identify *why* and *how* outcomes have come about. This is essential information in order to be able to attribute effects to program activities. It is also critical information in order to make decisions about future policies and programs, to assist in program improvement, and to be able to generalise what has happened from a single setting to elsewhere.

159. While monitoring assumes the validity of stated targets, evaluation can look at the continued appropriateness of the overall program or policy approach in its present form. Thus monitoring and evaluation information are complementary and both are essential for effective management and an orientation to performance. Monitoring may be sufficient, in some cases, to determine the presence of outputs. But some form of evaluation invariably would be required in order to identify outcomes, and to be able to provide for attribution. The following table summarises key differences between evaluation and monitoring.

Monitoring	Evaluation
<ul style="list-style-type: none"> • periodic; • assumes appropriateness of program, activities, indicators; • tracks progress against small number of targets or indicators; • usually quantitative; • uses data routinely gathered or readily obtainable; • cannot indicate causality; • difficult to use for impact assessment; • usually internal. 	<ul style="list-style-type: none"> • usually episodic; • can address a wide-range of potential questions about a policy, program or project; • can identify what has happened as a result of an intervention and provide guidance for future directions; • can address “how” & “why” questions; • wide-range of quantitative and qualitative research methods possible; • can use data from different sources; • can identify unintended as well as planned impacts and effects; • can involve internal, external or self evaluation.

160. There are a wide variety of potential evaluation approaches, ranging from the informal to the highly sophisticated, using a wide-range of both quantitative and qualitative methodological approaches. Evaluations can range from little or no cost to extremely expensive and time consuming undertakings. They can be participative in nature or carried out exclusively through external review. Evaluations can be carried out by programs directly (self evaluation), by internal evaluation units, or by external evaluators or a combination.

161. Evaluations can be carried at the end of a program in order to make judgements about its effectiveness or impact (known as *ex post* or summative). Evaluations can also be carried out *ex ante*, before a program is implemented, to assist in program planning. As well, they can be carried out while a program is running to assist with improvement (referred to as formative evaluations). Any form of evaluation has the potential to provide for understanding and for learning, about what works or not under given circumstances, and implications for future directions. Unlike monitoring, which is ongoing in nature, evaluations generally are episodic.

162. Nevertheless, as for example a recent presentation by Kristensen of OECD⁴² indicates, there recently has been some disillusionment with evaluation. There has been only limited use of evaluation where it has been carried out with no strategic purpose, such as when all programs are subject to comprehensive cyclical evaluation. While this approach is still used in some jurisdictions, Australia, for example, has recently backed away from this.

163. Evaluation can be most useful, and most cost effective, when it is used strategically, focusing on key questions of particular interest with action implications. Monitoring information can be invaluable in identifying specific questions for more intensive evaluation. The degree of sophistication in the research methodology should be based upon the intended use and the degree of confidence required. As a general rule, one should not aim for the most comprehensive research approach possible, but instead on how to provide the minimum amount of information needed to produce the necessary confidence to take required action. Evaluations that are highly focused on key questions of interest are likely to be more useful, as well as more cost effective, than large-scale research studies that attempt to be all-inclusive.

164. An integrated approach to monitoring and evaluation is likely to be most useful. Perhaps the best example of this is in Malaysia. Malaysia, in general, takes a comprehensive rather than a piecemeal approach to results-focused management and budgeting. Its results-based budgeting system (RBB) is focused on managing for results both at the output and outcome levels. It is also integrated into a range of complementary public sector management reform initiatives, all of which are intended to provide public sector managers the appropriate management tools to assist them improve their organisational performance as required under the results-based budgeting system. Malaysia has also carried out evaluations of its RBB system, and has made modifications and adjustments to its approach based upon information from the evaluations.



165. The basis of the performance management system is the use of a simple but powerful logic model called the ProLL (Program Logic & Linkages) Model.⁴³ The ProLL Model, shown above, was first developed by Dr. Arunaselam Rasappan for use in the Malaysian public sector. It is used both as a program planning tool as well as a tool to help identify critical evaluation questions right from the planning stages. This analytical process helps both program planners and program evaluators as it identifies and draws due attention to a number of key questions related to the “purpose-in-life” (PiL) of a particular program and its performance.

166. One of the strengths of the ProLL model is that it raises fundamental questions that are often taken for granted or ignored by program managers and planners. For example, it does not assume that program clients and their needs are clearly understood and specified. Similarly, the model requires detailed analysis of the clients' needs and problems before specifying the objectives. The treatment of the objectives is also very detailed, structured and practical to the PiL and desired results of the particular program and/or agency.

167. The ProLL model is very useful as a program planning tool and it is used both for planning new programs as well as for reassessing existing programs. The model is also useful as an evaluation tool when used in an evaluation mode. In the evaluation mode, the ProLL model is used to assist both program planners as well as evaluators to raise a number of critical evaluation questions pertaining to relevant components of the model. For example, the model raises evaluation questions as to whether the right clients were identified and their right needs met.

168. The model is extremely useful in identifying a series of evaluation questions that would normally be raised in an evaluation exercise. The advantage of the ProLL model however, is that it raises such key evaluation questions right from the beginning of the planning process. It thereby raises the relevant evaluation questions during the planning stages and assists program managers in working towards ensuring that these questions are addressed throughout the program.

169. This model can also be used to identify the different types of questions that can occur at various points, the types of performance measures, and what types of formative and summative evaluation information may be appropriate at each step. In this way, the program managers are able to better manage their program's performance and also systematically collect performance data on critical aspects as the program progresses along. The basis of this approach is both preventive and self-evaluation whereby program managers are able to carry out formative evaluation on an ongoing basis and thus ensure that the right results are achieved for the right target groups in the right manner.

170. As Section 5.5 observed, there has been very little evaluation of public sector reforms, including of results-focused approaches to management and budgeting. Malaysia appears to be one of the few exceptions. This may be one of the reasons for the limited focus on outcomes, as has been identified in Section 2.2.

6.2. *Whole-of-government approach*

171. The papers and presentations at the meeting reflected considerable interest in how to instil a cross-cutting or whole-of-government approach. There is increasing recognition that solutions to many of the issues that government needs to address (to give but two examples, poverty reduction and economic development) require action, and coordinated action, from across multiple program areas and agencies. Overly narrow approaches not only can be limited in their potential to be effective, but can "solve" one problem by creating problems elsewhere, in effect "shifting" problems and sometimes costs elsewhere. Vertical "silo" structures and thinking can be counter-productive.

172. Nearly all the country presentations acknowledged this as an important issue. Despite attempts to provide for more of a whole-of-government focus, such as joint activities across program areas and agencies, joint planning and joint target setting in some cases, this issue nevertheless remains problematic.

The pull of vertical silos remains strong. Despite exhortations to think globally, at the end of the day people still tend to think in terms of their own organisational units for their identity.

173. There are some practical reasons for this. Without counter-balancing horizontal mechanisms, program funding, both formal and informal recognition, promotion and other forms of rewards and sanctions mainly come through the vertical silo. The reality in some cases has been that managers and staff who work too hard on cross-cutting initiatives can be viewed as disloyal to their home base, which can harm their promotional prospects. Managers may be receiving conflicting messages. It is not clear to what extent a performance management focus can encourage, or conversely hinder, a whole-of-government perspective. Indeed, some of the countries specifically indicated that it can have a negative impact. For example, as one country put it:

Performance contracting ... has resulted in a higher priority being given to meeting agency objectives rather than promoting cross-cutting action [due to the] performance evaluation and reward system, which provides few incentives for agencies to focus on inter-departmental issues.

174. Silo thinking and action is not just confined to public sector organisations. It might be instructive to consider briefly how the private sector has attempted to engender more of a “cross-functional” approach. This has been done through two basic approaches:

- changes in management structure and practices;
- realignment of systems of recognition, including rewards and sanctions.

175. In the former case, the use of matrix management and cross-functional teams in order to address needs are commonplace. Some organisations go further, for example with extensive use of task forces and “SWOT” teams, where the members hardly have a home base.

176. There is also increasing recognition within the private sector that systems of rewards that are based upon performance to one’s home base can sabotage efforts at providing for a cross-functional approach. Performance reviews need to acknowledge the reality that priority to cross-functional activities often can hamper performance on other projects. There are a variety of ways in which this has been attempted.

177. **Real and substantial attention to cross-functional work.** Rather than giving lip service to a manager’s cross-functional efforts, effective performance reviews and decisions about compensation and promotion provide for *substantial* attention to cross-functional contributions, in order to give a real and tangible message that this is valued at least as highly as performance within one’s own organisational unit. One of the major criteria for rating performance can be the extent of cooperative work. If the cooperative activity is real and substantial, this generally would be expected to include concrete examples where cooperation has “hurt” performance in a silo perspective.

178. For example, development of staff, including loaning or promoting one’s *best* staff to other areas may be critical to the overall effectiveness of the entire organisation. But it can make little sense to a manager to do so, as losing one’s best staff can hamper the ability to meet one’s own targets within the silo. Where this has worked, in both the private and public sectors, there has been explicit recognition of

this situation, with compensation, rewards, and recognition provided accordingly. In some cases, a significant proportion of a manager's own performance rating has been based on the subsequent performance elsewhere in the organisation of their own former staff.

179. **Recognition based upon overall performance.** Where cross-functional contributions are important, recognition, including rewards and sanctions, can be based to a substantial extent not on individual performance or of that of a manager's own unit, but of overall performance. For example, strategies to improve the effectiveness of teams include rewarding members for how the team as a whole performs, irrespective of their own individual contribution. This also can be done on a larger scale, for example on the performance of an overall initiative involving the participation of different units. Along similar lines, where cooperation is primarily between two units, the performance assessment of a manager can be based to a large extent upon the performance of the *other* unit. This can provide genuine incentives for cooperation.

180. It should also be recognised that for many cross-cutting initiatives, there can be overall benefits to government/society, but with expenses incurred in one cost centre and the benefits realised elsewhere. As Pollitt⁴⁴ put it, in his OECD report: "Performance information about an entity may be distorted if the reported performances are, in part, being achieved in some other entity's budget." For example, providing for accessibility for people with disabilities, which frequently is a prerequisite to permit them to live and work in the community, can require expenditures by a Transport Ministry, but with benefits accruing in ministries dealing with social services, citizenship, and/or employment. Without explicit recognition of this, there is little incentive for managers to take action that, in their view, will incur costs to them but only benefit others.

181. **Cross-functional performance reviews.** Overall reviews or evaluations from a systems perspective have the potential to identify overall impact, as well as factors that are contributing or detracting from this.

182. While there are a number of potential strategies that can be attempted, there is no simple or obvious answer how best to facilitate a whole-of-government perspective, despite the recognition of its importance. This appears to be an issue worthy of further consideration, including sharing of experiences from across OECD countries.

6.3. *Implications of a results-oriented approach for external contracting*

183. In many OECD countries, including most of the Anglo-Saxon and Nordic countries, contracting or "outsourcing" with agencies outside government is widely used as a means of providing public services. The nature of services provided by non-governmental bodies can range widely, for example including: garbage collection, community-based services for people with disabilities, health services of various forms, adult education, even the operations of prisons and postal services.

184. The type of service provider also can vary, for example including:

- Arm's-length agencies (e.g. Crown corporations in the United Kingdom and Canada), that technically may be outside government but that are not completely independent and still may report to a minister in some way. Also falling under this category might be businesses, such as in Norway, in which the government owns a controlling share.
- For-profit independent private businesses.
- “Third sector” organisations, most typically not-for-profit organisations, sometimes with a heavy volunteer component (most frequently referred to as NGOs or non-governmental organisations).

185. The nature of relationship with government also can vary, for example ranging from being treated similarly to a supplier of office supplies, to a close partnership relationship.

186. There are a variety of reasons why some governments have decided to outsource certain services. In some cases, it is believed that routine services in particular can be provided more efficiently or cheaply from outside government. In some cases, the level of expertise that is needed may be found more readily in non-governmental organisations of some form. Community based or non-profit organisations may be seen as closer to the people and thus better able to gain the necessary trust and to be able to respond appropriately. Organisations based outside of government can be free of many of the constraints and rules applicable to services located directly within the government apparatus, and thus may have greater flexibility and be able to respond quicker to emerging needs. The influential book *Reinventing Government*⁴⁵ suggests that there are certain functions that government is best at providing, others in which the private sector is stronger, and still others which the third sector can best address.

187. Some outsourcing decisions seem to be based upon a belief that the best government is a small government and that government should divest itself as much as possible of direct service delivery. Thus a major driving force for outsourcing, at least in some jurisdictions, appears to be ideological rather than evidence based.

188. The major rationale for a results-focused approach to functions delivered outside government is essentially the same as for those provided directly by government, that is to shift the focus from input to results and on the actual value received in accordance with expenditures. Because the expenditures, and sometimes the outputs, of services delivered outside government may be more explicit than for direct government services, a results-focused approach also has the potential to provide for greater transparency.

189. In addition, a results-focused approach is sometimes viewed as a more objective, cost-effective way of purchasing services or outputs. When outputs can be specified with precision, this can provide a basis for indicating what government, and the public, gets for its money. This is thought to serve as a basis for making purchasing decisions, by providing for competitive bidding where the output is fixed and thus competitive bids can be assessed on the basis of price.

190. New Zealand perhaps has moved more than any other jurisdiction in the above direction. For example, it had separated government agencies into “purchasers” and “providers”, with extensive use of competitive purchase of outputs. More recently, however, it has been rethinking this approach, for example recombining government agencies that earlier had been separated. There is increasing awareness

of some of the limitations associated with focusing too much on outputs, as has been discussed earlier. These include, for example, concerns about how this approach can lead to actual outcomes and to larger government objectives, as well as dangers of goal displacement.

191. It is also worth noting that the private sector increasingly has been changing its own approach to the management of outsourcing. In particular, there is a movement away from using competitive tendering, at least as a basis for making decisions mainly or entirely on price, as a means of choosing amongst suppliers who are viewed as adversaries and essentially interchangeable.

192. In its place, a smaller number of long-term relationships are being established with whom are viewed and treated as partners rather than as suppliers of services or products. The lowest priced bid is not necessarily viewed as superior. This recognises that it is never possible to specify outputs exactly, and that low cost frequently comes at the price of quality. Thus when full costs and other factors are taken into consideration, a low-priced bid in the long-run may not result in the highest quality, or even the best value. Selection of suppliers is made upon the basis of a number of both subjective and objective criteria, with an emphasis on the long-term.

193. To be sure, performance and cost remain critical considerations. But this is approached in a cooperative, rather than an adversarial, fashion. For example, if there are problems, rather than immediately switching to a different supplier, assistance may be offered to attempt to solve them, as well as to identify ways in which performance can be increased and costs kept down, to the benefit of both parties. Frequently suppliers are given access to internal resources of the purchasing organisation, such as inclusion in training and development, on-line access to internal databases, research information, etc.

194. In addition, it is recognised that many suppliers may be closer to where the action is than the host company, including sometimes with more direct contact with customers. In some cases, suppliers may be better placed to identify what is and is not working well and why, and what types of changes or improvements may be called for. At the least, they can offer a different perspective that is recognised as valuable. Some organisations even provide for the active participation of key suppliers in their policy development and planning.

195. As a consequence, the lines between the internal and external organisations increasingly are becoming blurred, intentionally so. In other words, a performance *management*, rather than just performance *measurement* approach is being taken with outside organisations, with an emphasis on what can be done to improve current activities and to identify the need, where appropriate, to develop new approaches. It is recognised that it is important to facilitate the development of a results-oriented culture in external agencies, as well as directly amongst one's own employees.

196. The above approach, of course, is not unique to the private sector. Many governments also actively involve community-based service providers, for example, viewing them as partners, provide assistance as well as direction. These are expected to be long-term relationships, with withdrawal of funding viewed as a last-choice option.

197. **Implications for governments.** What are the implications of the above considerations for how governments can approach contracting? The following seem to emerge:⁴⁶

- Contracting does appear to be a viable option in a number of circumstances. But only certain functions are best located outside of government, and different types of non-governmental organisations are more appropriate for different types of functions.⁴⁷ Clear criteria should be developed, identifying the rationale for locating services outside of government (*e.g.* less expensive, more appropriate for reaching the target group, expertise required is located elsewhere), and the type of organisation most appropriate to provide the service.
- A broad, long-term perspective to outsourcing should be taken, at least with respect to key public services, where service providers are viewed as partners rather than as suppliers who can be replaced at a moment's notice. This approach implies a recognition that it generally is in everyone's interests for services to be as effective as possible, and that it is appropriate for government to provide assistance and support in this regard. It can provide for a focus on outcomes that can avoid an overly narrow focus on activities or outputs that can distort activities, and that can make provision for a responsive approach, including negotiation regarding changes in expected performance in order to be responsive to changing needs and context.
- Similarly, relationships with providers outside government should be viewed as an opportunity for two-way communications. For example, in many cases external service providers may be able to help bring government in some cases more in touch with the views of citizens.
- One should acknowledge, as identified in Section 4.2, that the use of outputs both for making decisions on competitive selection and for monitoring subsequent performance would be most appropriate for routine activities where the outputs can be easily established, which are taking place in a more or less stable environment so that agreed-upon targets are likely to remain appropriate, and where the impact of external factors on performance is expected to be minimal. Even in these circumstances, one should pay attention to the potential for goal displacement, in particular with respect to tradeoffs between meeting quantitative targets, and factors such as quality and the manner in which the services are provided.

One also should guard against an overly mechanistic approach to setting and using targets. While purchasing outputs can be appropriate in some cases, in other situations, however, this approach can impair rather than advance an outcome-focused approach.

- A performance management approach that includes both monitoring and evaluation can be very useful in managing services that are contracted out. This can provide a basis for determining what has been done and what has been achieved and to provide for accountability. In addition, it can provide objective evidence about the appropriateness of the activities that are provided, it can identify the reasons for problems, if any, and how they can be addressed, and suggest what both parties can do to maximise future performance.

198. It is important to bear in mind that the manner in which “outsourcing” is used to deliver key public services varies considerably across different countries and different systems of government. The practice is quite common in some countries, and rare in others. It is compatible with certain forms of government, less so with others. The nature of agencies entrusted with the delivery of public services can also vary, *e.g.* restricted to non-profit organisations in some situations. Each jurisdiction will have to identify what is most appropriate in its own circumstances, perhaps using the criteria suggested above as a starting point.

6.4. *Linking budget and results information*

199. How can – or should – information about performance be linked with budgetary considerations? There is keen interest in this topic among OECD Member countries, which was reflected in the discussions at the expert meeting. A recent OECD⁴⁸ report is specifically devoted to this topic.

200. Ideally, it only appears to make sense that results information would form part of the budgeting process. Indeed, this is implicit in the phrase “results-focused management and budgeting”, and is one of the goals of public sector reform and of performance management.⁴⁹ Good budgeting is one of the cornerstones of good governance, and it would seem that this would be strengthened when financial decisions are based upon solid evidence of what works.

201. Nevertheless, despite this ideal, as the Pollitt OECD report⁵⁰ discusses there are many difficulties to integration of financial and performance management. This report indicates that on those occasions where it has been tried, it rarely has worked as billed. This was confirmed by the presentations at the expert meeting, where there is considerable interest in at least some degree of integration, but where this is viewed as premature. Indeed, it is not clear that integration is even desirable or appropriate, at least not across the board.

202. There appear to be three basic reasons restricting the direct linkage of financial and performance management:

6.4.1. *Technical limitations in both performance and financial data*

203. As the Pollitt OECD report discusses, there can be a variety of technical difficulties with performance data, limiting their direct use for budgeting purposes, such as:

- As identified earlier, only certain types of programs, in particular those involving the delivery of tangible, standardised activities providing recurring products or services, readily lend themselves to performance measurement and thus to the calculation of reliable unit costs.
- Similarly, as has been discussed, outcome measures are difficult to construct for public sector programs, and this is a basic prerequisite to direct linkage to financial data.
- Many public sector activities, such as research, poverty reduction, environmental improvement, have outcomes that can only be achieved some time in the future, and certainly not on the same timescale as the budgetary cycle. Indeed, this is likely true of many public sector programs and interventions.

204. These, and similar issues, were also raised during the discussions at the expert meeting.

205. On the other side of the equation, good budgetary information is also required in order to be able to link expenditures with outcomes, or even with activities. This is recognised as a challenge, with fairly primitive cost recording still the norm in many countries. Most of the countries represented at the expert meeting are making improvements to the quality of financial information and accounting and budgeting processes their priority. For example, there was considerable attention at the expert meeting on how to move towards accrual accounting, activity-based and full-cost accounting.

206. Good cost information clearly is a basic component of effective financial management, and a prerequisite to integration with performance data. But it is important to bear in mind that this represents just one side of the equation. Some of the presentations seemed to imply that better accounting information and budgetary processes would enable them to be able to make budget decisions based upon performance. But as indicated above, valid and reliable performance information would also be necessary.

207. The following table which summarises key variables for integration is copied verbatim from the Pollitt OECD report:⁵¹

Integration would be easier in a context where:	Integration would be more difficult in a context where:
<ul style="list-style-type: none"> • Strategic target/objective setting is linked to resource allocation. • Global or output-based budgeting is in place. • Full cost activity accounting is in place. • The program in question consist of a set of tangible and measurable products or services. • Integration is being attempted at the levels of program priorities management and operational management. • The impact of a program can be seen soon after the services or products are delivered. • The results (outcomes) can be attributed to the program with high confidence (rather than there being reason to suspect that they were caused by other factors). 	<ul style="list-style-type: none"> • Historical incrementalism is the basis of resource planning and allocation. • Line item budgeting is in place. • The accounting entities do not match the units in which program activities are carried out and performance is measured. • The program consists of non-standardised, non-tangible, “ideal” services. • The effects of the program can only be detected in the long-term. • Even when “results” are detected, attribution directly to the program is uncertain.

6.4.2. *Performance data, by themselves, do not provide a sufficient basis to make financial decisions*

208. Even if technical limitations can be overcome, performance measurement data still provide only part of the story. At best, they are *indicators* of selective aspects of a program’s overall performance. And in order to make informed decisions, it is necessary to understand *why* the given levels of results were obtained. For example, poor performance could be a result of a program that is poorly conceptualised or is no longer relevant, in which case defunding might be an appropriate step. But there are many other potential reasons, with implications for funding. Poor performance could be a result of poor management, of the right program but delivered to the wrong client group or at the wrong location. Performance could be affected by factors beyond anyone’s control (e.g. an environmental disaster, changes in the macro- or micro-economic picture). Poor performance also could result from inappropriate staffing or funding, difficulties with measurement, or a range of other possible variables.

209. The United States report indicates that:

The causes of poor performance should be first understood before deciding what course of action to take. In certain circumstances, underfunding may have led to poor performance, and one remedy would be to increase the program funding level. A similar assessment of a well-performing program could conclude it is functioning optimally, and additional funds would yield little benefit or effect.

210. There is also the danger of performance information being taken out of context, if the link with budgeting is too mechanical. For example, in one jurisdiction, the Treasury cut the budget to a program involving inspectors checking for violations of fishing regulations, because performance was well below what had been targeted. However, the reason for this was because another government policy restricted usage of firearms, and in a well publicised labour action, the fishing inspectors refused to continue to apprehend potential violators. Responding to the changed circumstances, the program redirected the roles of the inspectors from apprehensions to a variety of educational activities. Yet Treasury, at least initially, did not take into account any reasons why performance may have been different than initially forecast.

211. Australia, during discussions at the expert meeting, warned of the danger of being overly mechanistic in setting and using targets and performance information. As the Australian representative indicated: “Mechanistic applications are probably the worst enemy of results-based management.”

6.4.3. *Inappropriate to make decisions about future directions just on the basis of past performance*

212. Performance data, by definition, is about what happened in the *past*. Decisions, however, including decisions about budgets, refer to what should be done in the *future*. Thus past performance can, at best, serve as only one factor to guide future directions.

213. As Kristensen, Groszyk and Bühler⁵² indicate: “Expectations must mirror reality. Performance measures, including outcome goals, are only one factor in policy- and resource decision-making.” There are a number of other factors besides technical information based upon “rational” analysis that also must be taken into account in making decisions about future directions. For example, decisions are usually taken based upon a consideration of alternatives. Even if a program is not working very effectively, what are the alternatives? For example, no one would seriously suggest shutting down an inefficient health care system unless there are better options.

214. Also, future decisions of necessity are based upon value judgements, which in a democratic system inevitably bring them into the political arena. Information about past performance does not necessarily say what future priorities should be. Decision-making often needs to taken into account the frequently divergent views of a range of stakeholders about what future actions are most appropriate. These may vary from that arising from rationale analysis within the government bureaucracy.

6.4.4. *How results and budget information could be linked*

215. The above discussion suggests that a one-to-one direct link between performance indicators and budget discussions generally is not possible or appropriate. Nevertheless, as Pollitt indicates in his OECD report, integration does not have to be viewed as all or nothing. It would not be responsible for a government to forsake completely any hope of linking resources and performance in at least some way. This report proposes an analytic framework and suggests some indicators of integration.

216. One should be careful to avoid throwing out the baby with the bath water. While a direct mechanistic link between results in terms of target realisation and budget is neither possible nor desirable, information about results clearly should represent *one* source of information for use in planning and decision-making, including decisions about future allocations. One should also remember, as Australia emphasised, that perfect information is not essential. Decisions are always made on the basis of imperfect information. Making available at least some information on performance may be able to add some greater degree of confidence to the decision-making process. At the least, it can be used to raise questions that should be considered or explored through other means.

217. The type of performance information available also can have a bearing on its usefulness. In particular, evaluations, depending upon how they are done, have the potential to generate more useful information than monitoring or tracking data. Evaluations that explore the reasons a program has worked or not, and what might make it more effective, can provide a future orientation and address many of the limitations of performance data for decision-making that were identified above. As well, evaluations have the potential to explore the pros and cons of alternative approaches, through a variety of techniques including *ex ante* evaluation and reviews or syntheses of evidence of effectiveness in a number of different settings.

218. In some cases, evaluations can identify the perspectives of different stakeholders, and sometimes arrive at a consensus. They may be able to anticipate the types of questions likely to arise, and provide at least some information to help with this. An integrated monitoring and evaluation approach, as suggested above, can help target evaluations where they are most likely to contribute to the policy development or decision-making process. For example, the European Union has placed considerable emphasis on developing a strong evaluation practice, in order to improve management and design of programs, to provide for accountability, and to contribute to decisions about resource allocations.

219. The priority to date among the countries represented at the expert meeting has been to provide information about results together with financial information, for example in budget documents and/or in annual reports. The underlying assumption is that this will increase transparency and provide at least for better informed decisions about appropriations. Even this has proved challenging at times, for example in coming up with the right format for including meaningful results information in the appropriate budgeting and accounting documents.

220. This approach seems to make sense. It might be further refined by thinking through *why* performance information should be included with budgetary information and how it is expected to be used. For example, is this mainly for transparency and accountability? Is it due to a belief that somehow performance information *should* be there, given the overall objectives of results-focused management and budgeting? Or is it expected to add to or influence the budgeting process in some way? Who would be expected to use the information? The most appropriate forms of reporting and presentation of information can vary depending upon the response to questions such as these.

221. It might also be a useful exercise to assess the *quality* of information that is presented. Boyle's⁵³ review of mandated annual reports produced by government departments and offices in Ireland is relevant in this respect. He found a number of deficiencies in these reports. For example, he found that reports are generally data deficient, mainly activity and output based, and for these and other reasons, impossible to assess progress against objectives and strategies. Reports focused almost exclusively on listing achievements, with little in the way of balanced discussion or consideration of the continued relevance of objectives and strategies.

222. Boyle's review also included consideration of international experiences on the quality of annual reports. He noted similar problems not just in Ireland, but in other jurisdictions as well, in particular by independent reviews in Canada and the United States. Drawing from international experience, Boyle discusses criteria for judging the quality of annual progress reports. These might serve as a useful starting point for OECD countries to evaluate the quality of their own results reporting.

6.5. *Results information and Parliaments*

223. Following from the above discussion, Parliaments constitute one potential group of decision-makers. Many results-focused reforms assume that results information *should* be before Parliaments, that this information *should* be of great interest to parliamentarians.

224. This, however, has resulted in not a little bit of frustration. Parliamentarians for the most part do not seem to be interested in results information. Some countries have held the view that this was mainly because the information delivered has been on inputs, processes and outputs and that things would change with the introduction of information on outcomes. Nevertheless, initial experience suggests that parliamentarians continue to ignore much of the abundant information actually available.

225. Based upon this experience, it might be useful to consider more specifically *why* it is considered desirable for parliamentarians to have information on results. Do parliamentarians really want, or need, this type of information? If so, in what forms?

226. To what extent do parliamentarians view it as their role or as a priority to oversee administrative operations, as opposed to identifying the need for new directions and developing appropriate legislation? Perhaps parliamentarians just want to be reassured that all is in order. If so, what level and types of information can do this and are realistic for busy parliamentarians to review?

227. Parliamentarians typically have considerable demands upon their time, with large piles of documents of various forms placed in front of them. Is it possible that reams of financial and performance data may have limited meaning for them, and be less interesting than other forms of "data" that they may be more accustomed to receiving, such as information from experts and constituents presented orally, as well as other documents that may be less dry than tables of performance data?

228. Assuming that there is at least some interest in information about performance, are there other forms of presentation that may obtain a better reception? Would more performance stories, such as suggested in Section 2.3, be more useful than raw data? A paradox is that too much aggregation can make what results nearly meaningless. But too much specificity can result in information overload. What is the most appropriate balance? Would there be any merit in a hyperlink approach, with enough information in summary form to enable busy readers to decide if they want to know more, with links to turn to for more detailed information.

229. As well, Kristensen *et al.*⁵⁴ have indicated what should be obvious, but can too easily be forgotten: parliamentarians may pay more attention to political considerations such as attracting media attention and re-election than to performance data. While “rational” management may be one objective, politicians of necessity have other interests to attend to as well, including responding to expressed desires and demands of the citizenry and of competing interest groups. This may be “irrational”, but another name for this is “democracy”.

230. It is difficult to answer questions such as those posed above in the abstract. Public sector reformers introducing results-focused management and budgeting approaches may wish to consider the extent to which their systems build upon a sufficient understanding of the dynamics of parliamentary politics. But this can represent dangerous ground, as it normally would be inappropriate for public servants to enter the political foray.

231. It should also be recognised that there is considerable difference from country to country in terms of government and parliamentary structures, roles, culture and traditions. Thus solutions to how best to present results information to parliamentarians may vary considerably.

6.6. *Taking into account the realities and complexities of public sector management*

232. Any results-focused approach to management and budgeting needs to acknowledge and reflect the realities and complexities of public sector management. Otherwise, it runs the risk of becoming a “rational” but largely irrelevant system and approach. As some of the participants at the expert meeting observed, such a situation can result in cynicism and detract from, rather than facilitate, actual results-oriented thinking and actions. Following are but a couple of considerations that should be borne in mind.

6.6.1. Managing in the context of multiple conflicting interests

233. It is useful to bear in mind that public sector management by its very nature is considerably more complex than in the private sector. In the private sector, the major imperative is maximising profits. It is possible for a chief executive to establish a clear vision and align people and processes behind this.

234. In the public sector, however, management constantly must operate in face of competing values, interests and goals, often in the face of intense public scrutiny. While “rational” management clearly is one key imperative, chief executives also need to respond to the direction from politicians which, as Kristensen *et al.*⁵⁵ have reminded us, may be different from what performance information would suggest.

235. Indeed, chief executives frequently need to respond to often conflicting guidance from the following:

- the Minister or head of their own agency;
- directions from the centre of government, at both the bureaucratic and political levels;
- perspectives and representations of key stakeholder groups, including sometimes from formal steering or advisory bodies.

236. In addition, public sector managers frequently must attend to rules and regulations that may serve a larger purpose, but frequently were not designed with what is in the best interests of implementing a particular initiative. Conversely, as we have seen, implementing whole-of-government approaches has proved challenging. Some countries in their presentations have acknowledged that performance measurement approaches in some instances have detracted rather than supported from this. And as the papers from Hungary and Sweden have reminded us, official goals and objectives can be political as much as realistic appraisals of what is possible to accomplish. This is one of many factors why meaningful assessment of performance needs to be broader than just identifying if pre-determined targets have been met.

237. Thus management in the public sector arena involves the art of the compromise as much as anything else.

6.6.2. *Complex adaptive systems and public sector management*

238. There has also been increasing attention to complexity theory, and in particular to characteristics of complex adaptive systems and implications for management. As Axelrod and Cohen⁵⁶ say: “In a world where many players are all adapting to each other and where the emerging future is extremely hard to predict, what actions should you take?” This situation can apply to much of what takes place within the public sector. Few significant public sector initiatives are independent, most of necessity must interact with a range of other initiatives and programs, government and otherwise, as well as with environmental considerations including social and economic aspects, and the perspectives of multiple stakeholders.

239. As Axelrod and Cohen say, under such circumstances one cannot *control* the future – but one can influence it. How can one manage for performance in such a situation of ambiguity, where the application of many traditional management and measurement techniques is unclear? Narrow reductionist approaches for assessing performance that attempt to oversimplify complex situations may be inappropriate. This appears to be an area that needs attention.

240. Mintzberg⁵⁷ has reminded us that good strategy (and this is just as true of strategy in the private sector as in the public sector) is as much emergent as predeterminant. One can never fix a strategy completely in advance, without responding to changed circumstances and what one learns along the way. This can have implications for how one approaches performance measurement.

6.6.3. *Importance of a balanced approach*

241. This paper has emphasised the importance of a balanced approach. For example, Section 2 has indicated the need across most countries for an increased focus on outcomes. This does not mean that inputs and outputs should be neglected, or that there might not be a need as well for better forms of cost data. Each country will need to assess what balance is most appropriate, which likely would vary depending upon the particular nature of each form of program activity. But given the current emphasis on inputs and, perhaps, on outputs, it would appear appropriate in most cases to shift the balance in order to devote more attention to outcomes.

242. Similarly, this paper has identified a number of potential problems and limitations with performance measurement data, particularly when these are used alone and out of context. Nevertheless, the paper has indicated ways in which this and other forms of monitoring can be very useful, provided that they are used appropriately as part of a broader strategy involving other forms of evaluation. Again, the most appropriate balance will need to be determined in each specific situation.

6.6.4. *Be modest and realistic*

243. Taking into account realities such as those discussed above can make a results-oriented approach all the more relevant and powerful. In particular, one should acknowledge that results information represents only one piece of what is needed for future planning and decision-making. One should bear in mind the aim of results-focused management and budgeting, which is not to follow a set of techniques, but instead to instil a results-oriented approach throughout government. As this paper has noted, public service administration is very complex. Attempts to reduce decisions about program effectiveness to simplistic formulas run the risk of missing important elements and providing misleading guidance.

244. Along these lines, as the Australian presentation emphasised, one need not aim for perfect information. Instead, one's goal should be the reduction of uncertainty. Instead of attempting to provide *the* definitive answer to complex questions, one should aim to add a bit more rationality and confidence to the policy-making and decision-making process within government.

245. As noted earlier, it is better to have imprecise answers to the important questions than to have precise answers to those that are unimportant. Performance information to the wrong question, however precise, is limited in utility. One should beware the danger of spurious precision, and recognise that outcome information generally is less specific, and more difficult to quantify, than data on activities and outputs. This principle can have implications for how one approaches monitoring and evaluation, as well as for how one interprets and uses the findings. For example, it can have implications for the choice of methods (*e.g.* quantitative *vs.* qualitative), as well as for the level of rigour.

246. One needs to set and to review progress towards intermediate objectives. But a long-term approach to performance management is needed. Most key government objectives, such as economic development, health promotion, or poverty reduction, are multi-faceted and long-term in nature, where the benefits of interventions not expected to be evidenced until some time after expenditures. This is also true with respect to achieving culture change.

NOTES

-
1. This paper is based upon a presentation to the OECD meeting: *Implementation Challenges in Results Focused Management and Budgeting*, 11-12 February 2002, Paris.
 2. OECD. *Evaluation and Outcome-focused Management and Budgeting*, presentation by Jens Kromann Kristensen to the TRANS-TALK Third Workshop: Improving Evaluation Practices in Transport, Brussels, 30 May – 1 June 2001.
 3. John Mayne. *Performance Measurement*. Evaluation Seminar 1. Evaluation Seminar Series. State of Victoria (Australia), Department. of Natural Resources and Environment, 2001. Available on the Internet through: www.nre.vic.gov.au.
 4. John Mayne. “Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly”. *Canadian Journal of Program Evaluation*. Vol. 16, No. 1, pp. 1-24, 2001.
 5. New Zealand Treasury. *Review of Evidence on Broad Outcome of Public Sector Management Regime*,” Treasury Working Paper 01/06 prepared by Murray Petrie and David Webber, 2001. Also available at: <http://www.treasury.govt.nz/workingpapers/2001/twp01-6.pdf>.
 6. This table is copied verbatim from Figure 2, Page 9 in Mayne, *op. cit.*, Note 4.
 7. OECD. *Outcomes-focused Budgeting — Case Studies*. PUMA/SBO(2001)5/ANN1. May 2001.
 8. Mayne, *op. cit.*, Note 4.
 9. *E.g.* see. Carol Weiss. *Evaluation: Methods of Studying Programs and Policies*. 2nd Ed. Sage, 1998.
 10. Mayne, *op. cit.*, Note 4.
 11. European Commission. *Evaluating EU Expenditure Programs: A Guide to Ex Post and Intermediate Evaluation*. 1997.
 12. OECD. *Knowledge Management; Learning-by-Comparing Experiences from Private Firms and Public Organisations*. Summary Record of the High Level Forum held in Copenhagen, 8-9 February 2001. PUMA/HRM(2001)3, CER1/CD(2001)2.
 13. One should note that there is an extensive literature on how to facilitate a utilisation-focused approach to monitoring and evaluation. *E.g.* see: Michael Quinn Patton. *Utilization-Focused Evaluation: The New Century Text*. Sage Publications, 1997.
 14. For but a couple of examples, see Patton, *ibid.*, and Robert F. Freeland. “When Organizational Messiness Works”. *Harvard Business Review*, Vol., 80, No. 5, pp. 24-25.
 15. OECD. *Public Sector Modernisation: A Ten Year Perspective*. Background document for the 24th session of the Public Management Committee, Château de la Muette, Paris, 5-6 November 2001.

16. Scott Bayley. “Measuring Customer Satisfaction: Comparing Traditional and Latent Trait Approaches Using the Auditor-General’s Client Survey”. *Evaluation Journal of Australasia*, Vol. 1, No. 1, pp. 8-17, March 2001.
17. Peter Senge. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Century, 1990.
18. OECD. *Engaging Citizens in Policy-Making: Information, Consultation and Public Participation*. OECD Public Management Policy Brief, PUMA Policy Brief No. 10, July 2001.
19. *Ibid.*
20. Christopher Pollitt. *Integrating Financial Management and Performance Management*. OECD report: PUMA/SMO(99)4/FINAL, 1999. Also available at: <http://www.oecd.org/puma>.
21. Indeed, performance measurement itself is often considered to be part of a command-and-control management style. *E.g.* see Henry Mintzberg. “Managing Government, Governing Management”. *Harvard Business Review*, Vol., 74, No. 3, pp. 75-83, 1996; Ian Sanderson. “Evaluation and Social Change: Coping with Complexity and Self-Organisation”. Presentation to the UK Evaluation Society Annual Conference, Belfast, December 2001.
22. *E.g.* see Burt Perrin. “How to – and How Not to – Evaluate Innovation”. *Evaluation*, Vol. 8, No. 1, pp. 13-28, 2002.
23. Mayne, *op. cit.*, Note 4.
24. *E.g.* Burt Perrin. “Effective Use and Misuse of Performance Measurement”. *American Journal of Evaluation*, Vol. 19, No. 3, pp. 367-369, 1998; Perrin. “Performance Measurement: Does the Reality Match the Rhetoric?”. *American Journal of Evaluation*, Vol. 20, No. 1, pp. 101-114, 1999; Jerome Winston. “Understanding Performance Measurement: A Response to Perrin”. *American Journal of Evaluation*, Vol. 20, No. 1, pp. 95-100, 1999; Ian C. Davies. “Evaluation and Performance Management in Government”. *Evaluation*, Vol. 5, No. 2, pp. 150-159, 1999; Jennifer C. Greene. “The Inequality of Performance Measurements”. *Evaluation*, Vol. 5, No. 2, pp. 160-172, 1999; A. B. Blalock, (1999) “Evaluation Research and the Performance Management Movement: From Estrangement to Useful Integration”. *Evaluation*, Vol. 5, No. 2, pp. 117-149, 1999; Henry Mintzberg, *The Rise and Fall of Strategic Planning*. New York: The Free Press, 1994; and Mintzberg. “Managing Government, Governing Management”. *Harvard Business Review*, Vol., 74, No. 3, pp. 75-83, 1996; Kathryn E. Newcomer. *Using Performance Measurement to Improve Public and Nonprofit Programs*. New Directions for Evaluation, No. 75. San Francisco: Jossey-Bass; Lena Lindgren. “The Non-profit Sector Meets the Performance-Management Movement: A Program-theory Approach”. *Evaluation*, Vol. 7, No. 3, pp. 285-303, 2001.
25. *E.g.* see the sources mentioned in the above note.
26. Cited in: Midge F. Smith. *Evaluability Assessment: A Practical approach*. Kluwer, 1989.
27. “China’s Whopper of a Fish Tale”. *International Herald Tribune*. 1 December 2001. Also see: “The Seas Need Care”. *The New York Times*. 9 January 2002.
28. Michael C. Jensen. “Corporate Budgeting is Broken – Let’s Fix It”. *Harvard Business Review*. Vol. 29, No. 10, pp. 94-101, November 2001.
29. *The Economist*. “Why Honesty is the Best Policy: Corporate Deceit is a Slippery Slope.” 9 Vol. 362, No. 8263, March 2002.

-
30. Henry Mintzberg. “Managing Government, Governing Management”. *Harvard Business Review*, Vol. 74, No. 3, pp. 75-83, 1996.
31. Harris Collingwood. “The Earnings Game: Everyone Plays, Nobody Wins”. *Harvard Business Review*. Vol. 79, No. 6, pp. 65-74, June 2001.
32. *E.g.* see Jensen, *op. cit.*, Note 28.
33. Richard Boyle. *The Use of Rewards in Civil Service Management*. CPMR Discussion Paper 5. Committee for Public Management Research, Ireland Department of Finance. 2001. Available for downloading from the Internet at; www.irlgov.ie/cpmr.
34. Mintzberg, *op. cit.*, Note 30.
35. Pollitt, *op. cit.*, Note 20.
36. *E.g.* Christopher Pollitt and Geert Bouckaert. *Public Management Reform: A Comparative Analysis*. Oxford University Press, 2000.
37. OECD (Knowledge Management), *op. cit.*, Note 12, OECD (Public Sector Modernisation), *op. Cit.*, Note 15.
38. Rickard Boyle. *A Review of Annual Progress Reports*. CPMR Discussion Paper 18. Committee for Public Management Research, Ireland Department of Finance. 2001. Available for downloading from the Internet at; www.irlgov.ie/cpmr.
39. Michael Power. *The Audit Society: Rituals of Verification*. Oxford, 1997.
40. *Op. cit.*, Note 11.
41. UNICEF. *A UNICEF Guide for Monitoring and Evaluation*. Evaluation Office, UNICEF, New York, 1991.
42. OECD, *op. cit.*, Note 2.
43. Based upon information supplied by Dr Arunaselam Rasappan. See also: “Results Oriented Budgeting Reforms: Malaysia & Singapore”. Presentation to the World Bank, July 1999; Rasappan. “How to Budget for Results”. *Khidmat*, pp. 18-20, September 1995.
44. Pollitt, *op. cit.*, Note 35.
45. David Osborne and Ted Gaebler. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. New York and London: Plume (Penguin): 1993.
46. Also see: OECD. *Policies to Enhance Efficiency in Publicly Funded Services: Issues and Experiences among OECD Countries*. Economics Department, Economic Policy Committee. ECO/CPE/WP1(2002)6, March 2002.
47. *E.g.* see Osborne and Gaebler, *op. cit.*, Note 45.
48. Pollitt, *op. Cit.*, Note 44.

49. E.g. see OECD. *Governance in Transition: Public Management Reforms in OECD Countries*. PUMA. 1995.
50. Pollitt, *op. cit.*, Note 44.
51. Pollitt, *op. cit.*, Note 44. Table 2, p. 23. Copied verbatim.
52. Jens Kromann Kristensen, Walter S. Groszyk, and Bernd Bühler. “Outcome-focused Management and Budgeting”. *OECD Journal on Budgeting*. Vol. 1, No. 4, 2002.
53. Boyle, *op. cit.*, Note 38.
54. *Op. cit.*, Note 52.
55. *Op. cit.*, Note 52.
56. Robert Axelrod and Michael D. Cohen. *Harnessing Complexity: Organizational Implications of a Scientific Frontier*. Basic Books, 2000.
57. E.g. Henry Mintzberg, *The Rise and Fall of Strategic Planning*. New York: The Free Press, 1994; Henry Mintzberg, Bruce Ahlstrand, and Joseph Lampel. *Strategy Safari*. Prentice Hall, 1998.