



EVALUATION CAPACITY DEVELOPMENT

ECD WORKING PAPER SERIES NUMBER 4: MAY 1998

The Development of Australia's Evaluation System

Keith Mackay



The Development of Australia's Evaluation System

Keith Mackay¹

Contents	Page
Executive summary	1
I. Genesis and stages of development	2
II. The development of evaluation capacity: evidence	6
III. Success factors and impediments— what has/has not worked, and why	10
IV. Current developments and prospects	16
V. Conclusions from the Australian experience	18
VI. Some implications for the World Bank	19
Bibliography	21
Endnotes	23

Executive Summary

The Australian federal government has given a high priority to ensuring that evaluations of its programs are conducted, and the findings used. The approach, which has been adopted to develop evaluation capacity, has entailed a combination of formal requirements for evaluation plus their strong advocacy by a powerful, central department (the Department of Finance). This has enabled evaluation to be linked both to budget decision-making and to the on-going management of government programs.

As a result, there is clear evidence of a high level of evaluative activity, and that evaluations are actually used to assist Cabinet's decision-making and prioritization in the budget, and to support internal program management within line departments.

Each country is unique, and the Australian success has been supported by the existence and ongoing development of a sound public service infrastructure — including a high level of institutional and human capacity. Nevertheless, many lessons from the Australian case are highly relevant to the development of evaluation capacity in developing countries.

Many success factors are identified in this paper. The key factors have included:

- the creation of an explicit, whole-of-government evaluation strategy;
- the existence of a powerful central department which has been a committed champion of evaluation, and which has continually identified and created new opportunities for influence and development;
- sustained commitment and support of the evaluation strategy over a decade; and
- implementation of related public sector management reforms which have given considerable autonomy to line managers and which emphasize bottom-line results — these reforms have provided incentives to conduct and use evaluation findings.

More recently, the Australian evaluation system has evolved from one of tight, formal controls and requirements to a more voluntary, principles-based approach. In this new environment it is hoped that the strong pressures for line departments to achieve and to demonstrate high levels of performance will be increased, and that the existing evaluative culture and infrastructure will be strengthened. The latest reforms include, for example, sharper accountability and performance standards for managers, the widespread application of competitive tendering and contracting of public service activities, and the setting and reporting of explicit customer service standards. The latest wave of public sector reforms promise a closer integration of performance measurement — including evaluation — into performance management and into governance more broadly.

I. Genesis and Stages of Development

An understanding of the development of evaluation capacity in the Australian government is important because of the insights it provides for other countries. Evaluation capacity development (ECD) did not progress in a linear, logical sequence in Australia; it grew opportunistically in response to prevailing imperatives. The twists and turns of these developments mean that much experience has been amassed concerning what works, what does not, and why. The objective of this paper is to share these insights.

As with other developed countries, evaluation is not a new phenomenon in Australia. Cost-benefit analysis — the appraisal of investment projects — has been part of the scene for many decades. It has been conducted by various government economic research bureaus, and by other advisory bodies and line departments concerned with infrastructure and agriculture investments.

Formal program evaluation has a more recent origin. It too has been conducted by specialist research bureaus attached to line departments, and by special policy review task forces, focusing on areas such as labor market programs and social welfare policy. But as a discrete and on-going focus of government activity its heyday started in the 1980s.

The 1983 genesis of public sector reforms

The election of a reformist Labor government in 1983 provided an environment favorable to evaluation. The new government was determined to improve the performance of the public sector while at the same time achieving tight public expenditure restraint through the annual budgetary process.

A series of public sector management reforms was implemented in the first several years of the new government. One aspect of these reforms was the desire to ‘let the managers manage’ by devolution of powers and responsibilities — reflecting the philosophy that public sector managers would be strongly encouraged to improve their performance if they were provided greater autonomy and the potential to manage their departments with fewer central agency controls and less interference. Another aspect of the reforms was ‘making the managers manage’, and this thread was more directly linked to powerful budgetary pressures.

The tangible changes included:

- substantial autonomy for departments in their spending of administrative expenses (including salaries), but with these administrative expenses being strictly cash-limited;
- greater surety about future resource availability to departmental managers via a system of three-year forward estimates of administrative and all other program expenses; and
- a major reduction in the number of departments through

amalgamation, to achieve less balkanized policy advice and to encourage the internal reallocation of resources through portfolio budgeting.

A related set of principles was embodied in the Financial Management Improvement Program, which included program management and budgeting. These principles emphasized the importance for departmental managers in ensuring that program objectives were realistic, to help guide managers and staff. The principles also encompassed a focus on the efficiency and effectiveness of programs — i.e., on program performance — through sound management practices, the collection of performance information and the regular undertaking of program evaluation. Guidance material which extolled the virtues of these principles and clarified concepts was issued by the Department of Finance (DoF) and the then Public Service Board, another central agency.

The public sector reform initiatives were introduced in a stringent budgetary climate. Macroeconomic concerns provided sufficient motivation to induce the government to reduce the share of federal government outlays in GDP from 30% in 1984-85 to just over 23% in 1989-90. The impact of these cuts on government programs was even greater than these raw statistics would suggest, because the government was also determined to increase the real level of welfare payments to the most disadvantaged in society. It achieved this by means of tight targeting of benefits via means-testing, which entailed substantial reductions in ‘middle-class welfare’.

This package of financial management and budgetary reforms was substantive and wide-ranging. For several years it placed the framework of public sector management in Australia at the forefront of developed nations.

These initiatives helped to create an enabling environment which encouraged performance management. However, while the reform initiatives were necessary they were not sufficient to ensure that evaluation and other types of performance measurement became accepted as desirable and routine activities.

DoF was a major architect of many of these reforms, reflecting its role as budget coordinator and overseer of the spending of other departments. DoF was keen to get out of the detail of spending issues, where it was often bogged down in minor spending bids and disputes with departments. Its concern with budget spending encompassed both a priority on cutting government outlays, and finding ways to make spending more efficient and effective.

Growing focus on evaluation

The concern with ‘value-for-money’ — that is, the efficiency and effectiveness of public expenditure — helped to lay the groundwork for DoF’s provision of advice to departments on the evaluation of their programs. DoF began this formally with the publication of an evaluation handbook in 1986.

At around the same time there was growing disquiet in DoF and other central agencies about the lack of any real progress made by line departments in managing their performance. In early 1987 the Minister for Finance took secured Cabinet’s agreement to a formal requirement that all new policy proposals for Cabinet consideration should include a statement of objectives and performance measures, as well as proposed arrangements for future evaluation. Line Ministers and their departments were required to develop plans for the systematic and comprehensive monitoring and evaluation of the performance of their programs, and for the reporting of these reviews to government. DoF was to be kept informed of departments’ evaluation plans.

DoF augmented its earlier advisory panel through providing additional guidance material in 1987, and via presenting a basic evaluation training course. By early 1988, however, the evaluation plans prepared by departments could best be described as patchy (many were poor), and it had become apparent that a more

fundamental examination of evaluation practices in departments was warranted. Thus DoF undertook a diagnostic study reviewing departments’ evaluation progress and the overall health of evaluation activities in the public service. The study found:

- a lack of integration of evaluation into corporate and financial decision-making;
- that evaluations tended to focus on efficiency and process issues rather than on the more fundamental question of overall program effectiveness — i.e. whether or not programs were actually meeting their objectives;
- a poor level of evaluation skills and analytical capacity; and
- that the role of central departments, especially DoF, was unclear.

The formal evaluation strategy

This diagnostic study laid the groundwork for a major submission from the Minister for Finance to Cabinet in late 1988 seeking — and securing — its agreement to a formal, ongoing evaluation strategy for all departments. A key and continuing principle underlying this strategy was that ‘the primary responsibility for determining evaluation priorities, preparation of evaluation plans and conduct of evaluations rests’ (with line departments).

Cabinet’s agreement to the evaluation strategy was expressed in a formal cabinet decision. For the federal government and its public servants, such decisions virtually have the force of a law. Indeed, efficiency audits conducted by the Australian National Audit Office (ANAO) often focus on the extent to which cabinet decisions have been complied with. Thus expressing public sector reform initiatives as cabinet decisions has been a way of encouraging the oversight involvement of the ANAO.

The evaluation strategy had three main objectives. It encouraged program managers within departments to use evaluation for the improvement of their programs’ performance. It also provided fundamental information about program performance to aid Cabinet’s decision-making and prioritization, particularly in the annual budget process when a large number of competing proposals are advocated by individual Ministers. Lastly, the strategy aimed to strengthen accountability in a devolved environment by providing formal evidence of program managers’ oversight and management of program resources. The emphasis was on trans-

parency, which is of interest to the parliament, particularly in the senate's processes of budget scrutiny and approval.

The evaluation strategy has provided the framework and driving force underlying the progress with ECD in the Australian government since that time. Its key components comprised four formal requirements. These were:

- that every program be evaluated every 3-5 years;
- that each portfolio (comprising a line department plus outrider agencies) prepare an annual portfolio evaluation plan (PEP), with a 3-year forward coverage, and submit it to DoF. PEPs comprise major program evaluations with substantial resource or policy implications;
- that Ministers' new policy proposals include a statement of proposed arrangements for future evaluation; and
- that completed evaluation reports should normally be published, unless there exist important policy sensitivity, national security or commercial-in-confidence considerations, and that the budget documentation which departments table in parliament each year should also report major evaluation findings.

A crucial aspect of this evaluation strategy was the role which Cabinet agreed to assign to DoF. Cabinet expressed its expectation that DoF would have the opportunity to make an input to PEPs and to the terms of reference of individual evaluations to ensure that these were consistent with government-wide policies and priorities, and that DoF would be available to participate directly in selected evaluations, subject to negotiation between DoF and the line department (or between their Ministers if a dispute arose). DoF was also to provide detailed advice and handbooks on evaluation methodology, and on management information systems, and take the lead in identifying and sharing best practice.

A key to the public sector reforms which DoF had advocated was the need for central departments to get away from detail and out of a control mentality. While the leadership of DoF were true believers in the utility of evaluation, they had advocated a more hands-on involvement for DoF with some hesitation, because it represented an activist and closely monitoring approach. DoF advocated a powerful role for itself to Cabinet only after there was clear and strong evidence of the failure of line departments to live up to the rhetoric of program management and budgeting which departments had also espoused. It would also probably be fair to say that

the immediate focus of the evaluation strategy was to ensure that good evaluations of the right programs were available, rather than the creation of a performance-oriented culture. The latter focus has emerged as the main area of emphasis only relatively recently (see section IV).

While the evaluation strategy was still at the proposal stage, most line departments stated their acceptance of the utility of evaluation but expressed real concern about 'intrusive' DoF involvement. This stated acceptance, at odds with the actual reality of management in most line departments, was a feature of other performance measurement and performance management initiatives introduced in later years.

Later developments

The next milestones for ECD in Australia were two reports, from a parliamentary committee in 1990 and from the ANAO in 1991. These reports acknowledged the substantial effort being devoted to the planning and conduct of evaluation, but argued for renewed efforts. In particular, they noted the variation in the extent of evaluation activity in different departments. They criticized some departments for their poor choice of programs to evaluate, and for the focus of, or issues covered by, their evaluations — particularly an insufficient focus on effectiveness issues. They argued that DoF should be more active in encouraging departments to plan and undertake evaluations.

These reports were soon followed by the creation in 1991 of a separate branch within DoF, responsible for providing evaluation advice, support, training and encouragement to other departments (and also within DoF itself). This branch which had nine evaluators, acted as a focal point and catalyst for evaluation throughout the Australian public service. The branch was co-located with two other branches, responsible for overall coordination and management of the government's annual budget process, and for public sector management reforms more generally.

Evaluation in the Australian government — as measured by the extent of evaluation planning, conduct and use — had achieved a healthy and vigorous state by the mid-1990s.

However, by that time DoF was concerned about departments' poor progress in articulating clear and achievable objectives for their

programs, and in collecting and reporting meaningful performance information. These concerns were confirmed by two reviews of departments' annual reports and of their budget documentation which DoF commissioned. This might appear paradoxical, because evaluation is at the more difficult end of the performance measurement spectrum, and was generally being done well, yet the setting of program objectives and the collection of frequent performance information, at the easier end of the spectrum, were being done poorly. The situation reflects the emphasis and success of the evaluation strategy in encouraging and mandating evaluation, with much less emphasis having been placed on ongoing performance monitoring.

Thus in 1995 DoF secured Cabinet's agreement to a rolling series of comprehensive reviews (staggered over 3 years) of the program objectives and performance information of all programs in all departments. These reviews are being conducted jointly by DoF and each line department, with the results being reported to their respective Ministers and to Cabinet as a whole. The reviews focus in part on analyzing the existing situation, but attach more importance to identifying ways in which objectives and performance information could be improved, and in mapping out and committing to a plan of action to achieve these improvements.

This illustrates, once again, that public sector management initiatives often require a push from the center to make them happen in many departments.

The election of a Liberal/National Party government in 1996 has led to an emphasis on cutting bureaucracy, red tape and the large number of formal reporting requirements. The new government required that existing programs be comprehensively reviewed by Ministers and their departments to establish whether they should continue, or be abolished or devolved to another level of government. For those programs which were to continue to be delivered at the federal level, Cabinet expected that a competitive tendering and contracting process be undertaken wherever possible.

One issue which had emerged with the formal evaluation requirements in recent years was that their concern with bureaucratic process was no longer appropriate — the length of some portfolio evaluation plans, for example, had grown from a recommended 20 or 30 pages to over 120 pages. A consensus had emerged within the bureaucracy that while it was important to have evaluation findings

available to assist decision-making by program managers and by Cabinet, detailed and elegantly-worded plans were not necessary to achieve that end.

A related and powerful strand of thinking held that departments should not be encumbered by excessive controls on their internal activities, as long as departmental heads and senior executives are responsible for performance, and that responsibility is reflected in their employment contracts. This is essentially a 'let the managers manage' philosophy, and is analogous to the one adopted in the early 1980s. A difference, however, is the greater scrutiny by Ministers of the performance of their department heads — as exemplified in the more widespread use of employment contracts — plus the apparent readiness of Ministers to remove department heads where their performance is judged to be wanting. And perhaps the most important difference is the progress in the intervening years in establishing an evaluation culture, and in establishing departmental infrastructures to support performance measurement and management.

This evolution in philosophy led in late 1997 to a further development in the evaluation strategy, into a principles-based, performance management framework. This approach was accepted by Cabinet, and is now government policy. Such a Cabinet-endorsed, principles-based approach provides guidance to the heads of line departments by emphasizing the good practice features of performance management and measurement (the latter includes evaluation and ongoing performance monitoring). It reflects the strong expectation that CEOs and senior executives will continue to plan, conduct and use evaluation, and so it implicitly takes the progress achieved to date as a given. These issues go to the heart of the question whether central controls should be tight or loose. A fuller discussion of the principles-based approach and these other issues is presented in section IV.

II. The Development of Evaluation Capacity: Evidence

It is important to have a realistic understanding of the extent of ECD in Australia. National reviews of progress with evaluation in developed countries often present too rosy a picture, particularly when external reviewers confuse rhetoric with reality. Checklists of evaluation activities undertaken do not necessarily translate into the widespread conduct of evaluation, nor into quality and rigor in evaluation, nor into its actual use in program management and government decision-making.

The true extent of ECD in Australia can be assessed by considering the planning, conduct, quality and use of evaluation.

Evaluation planning

All government departments have prepared portfolio evaluation plans since 1987-88. These were intended to comprise the major evaluations in each department and its outrider agencies — in recent years about 160 of these evaluations had been underway at any given time. Most of these evaluations were major, in that the programs had significant policy or spending implications, although a significant minority particularly for the smaller departments, were only of minor programs or of the efficiency aspects of large programs. (The plan guidelines issued by DoF recommended that the main focus of these evaluations be on issues of program effectiveness. Departments were separately encouraged to plan and to undertake more minor evaluations for their own internal management purposes.)

Line departments themselves decided which programs should be included in the plans for evaluation, and which issues the evaluation terms of reference would cover. However, DoF would usually endeavor to influence departments' choice of evaluation priorities by making direct suggestions. In doing so DoF would attempt both to anticipate and to create the information needs of Cabinet. Where DoF has had difficulty in persuading departments, it has sometimes approached Cabinet directly to seek its endorsement of particular evaluation suggestions and of detailed terms of reference; Cabinet almost invariably accepts DoF's suggestions.

The Cabinet-endorsed, formal requirement under the evaluation strategy that portfolio evaluation plans be prepared and submitted to DoF certainly provided a powerful incentive to line departments to prepare plans and to take them seriously. Another influential factor was the issuing by DoF of formal guidelines to departments on the desirable content of these plans, together with follow-up monitoring and reminders to departments about the need for the

plans. The evaluation branch of DoF conducted internal reviews of the content and coverage of these evaluation plans, and provided feedback and prompting to departments as well as by identifying good practice examples. In seven efficiency audits and two 'better practice' guides on program evaluation and performance information, the ANAO has also repeatedly reminded departments about the importance of systematically planning their evaluation activity.

The formal requirement that all programs be evaluated every 3-5 years was also influential in creating a climate in which evaluation is the norm rather than the exception. The concept of regular, comprehensive coverage of programs also encouraged a planned, staged approach to evaluation. This formal requirement should not be accepted at face value, however. It is very seldom that all aspects of a program² are included in any single evaluation. Instead, it is usual that an evaluation will focus only on certain key problems or aspects of a program. The challenge is to ensure that these difficult issues are actually evaluated, and this is where DoF has been active via persuasion and direct involvement in individual evaluations.³

Conduct of evaluation

Most departments have chosen to set up evaluation units to coordinate their formal evaluation planning. At their smallest, these units comprise two or three individuals. In some departments, such as employment, a separate branch of 20-25 staff is responsible for evaluation planning, provision of advice on evaluation methodology, participation in steering committees, and the conduct of major evaluations, particularly in the area of labor market programs (but typically not of education programs, which also comprise a substantial proportion of the department).

There is no standard approach in departments as to how evaluations will be conducted — this is viewed as a line management

decision. Some evaluations involve a wide array of external and internal stakeholders, either by participation in an evaluation steering committee, or less commonly by participation in the actual evaluation team. Some evaluations are conducted by a central evaluation unit, but it is more common for responsibility for the conduct of evaluations to rest with the line program area. These line areas would be responsible to the top management of the department for the quality and rigor of the evaluation.

For the more important evaluations — those listed in portfolio evaluation plans — some external involvement would be typical, via provision of suggestions and comments on the terms of reference and proposed evaluation methodology, participation in the steering committee, and provision of comments on drafts of the evaluation report. But there is no standard approach to this external involvement — it would be determined by the willingness of the line department to involve outsiders, and also by the interest and availability of outsiders such as central agencies. For programs with major resource or policy implications, DoF would usually be very keen to be involved, and would apply pressure to ensure its participation.

A recent ANAO survey found that, for evaluations conducted over the period 1995-1997: about half examined the delivery of products or services to external clients, and a further 30% were associated with matters internal to the department. One third of the evaluations examined the appropriateness of new or established programs, and 15% were directed at the development of policy advice for the government.⁴

The large number of evaluations in progress, and the fact that over 530 evaluation reports have been published over the last four years or so, attest to the existence of extensive evaluation activity in the Australian government. This has provided a growing library of evaluation findings.

DoF publishes a register of published evaluation reports, and this helps to monitor the progress of individual departments' activities. More importantly, it helps to share evaluation practices and methods among departments and this provides some quality assurance because the public availability of these reports exposes them to peer scrutiny. A recent survey of all departments by the ANAO found that 75% of evaluations conducted in 1995 and 1996 were released to the public or were available on request.⁵

Evaluation quality

Quality of evaluation reports is a more difficult dimension to measure. The rigor of program evaluations depends on the expertise and objectivity of the evaluators. A recent assessment of the quality of a small sample of evaluation reports was commissioned by the ANAO. It found that over a third of a sample of evaluation reports suffered from methodological weaknesses. It is certainly the case that some published evaluations are of low quality, and the suspicion is that some of these are produced for self-serving purposes, such as to provide a justification for the retention or expansion of the program.

DoF's own experience of evaluations is that their quality can vary enormously. The extent to which this should be a matter of concern is another matter: the issue to consider here is the intended uses of evaluations. If the intended audience of an evaluation is Cabinet (to aid its decision-making) or the parliament (for accountability purposes) then a poor quality or misleading evaluation gives cause for serious concern. DoF has certainly been willing to provide Cabinet with a dissenting view on the quality of an evaluation in those cases where an evaluation is used by a line department to attempt to influence a Cabinet debate. (Line departments would typically try hard to avoid such disagreements, which would be virtually guaranteed to attract the ire and condemnation of Cabinet.)

Where line departments allow poor-quality evaluations to be conducted (and these evaluations are intended for internal program management purposes) however, then there is perhaps an element of caveat emptor. Thus the extent of evaluation quality assurance or quality control could be regarded as an issue for departments' internal management to address.

A commonly-asked question is how evaluation quality can be assured, and what the role of DoF is or should be in guaranteeing quality. In past years, parliamentary committees and the ANAO have argued that DoF should take a strong role as an independent check on departments. DoF has preferred to seek to participate directly in certain major evaluations, usually via steering committee membership, thus ensuring that evaluations address the difficult questions and do so in a rigorous manner. But it would be a very resource-intensive activity to undertake detailed reviews of all evaluations, and it would also be inconsistent with the devolu-

tionary reform philosophy for DoF to do so.

The ANAO has consistently argued that departments should set up central oversight procedures to achieve quality assurance of evaluations conducted by line areas within the department. There is certainly evidence from those few departments which have followed this approach, that it is an effective means of bringing to bear needed evaluation skills and expertise, and of ensuring evaluation quality.

Use of evaluation

A bottom-line issue is the extent to which evaluation results are actually used. If their use is patchy or poor then there really is little point in conducting evaluations. The large volume of evaluation activity in itself provides some reassurance that evaluation findings are being used — in an era of very tightly limited administrative expenses, departments would not bother to conduct evaluations unless they were going to be used.⁶ (And DoF would not bother to advocate and work to influence the evaluation agenda unless it perceived high potential value in their findings.)

There are differences in the perspectives of a line department and a central agency, of course, and these influence the types of evaluation which are conducted and their probable uses. Line departments have traditionally been focused on evaluations concerned with program management and improvement, whereas the primary focus of central agencies is on overall assessments of the worth — the cost-effectiveness — of the program. But this distinction has become considerably blurred in recent years with the initiative of portfolio budgeting — this provides line Ministers and their departments much greater say in making decisions about portfolio spending and program priorities (see section III). Portfolio budgeting has encouraged line departments to focus more on cost-effectiveness.

A recent survey of departments by the ANAO found that the impact or use of evaluations was most significant with respect to improvements in operational efficiency, and to a lesser extent to resource allocation decisions and the design of service quality improvements for the benefit of clients.⁷

There is clear evidence that evaluations have been used intensively in past years in the budget process. DoF has conducted several

A 1994 review of evaluation activities in one of the largest departments, the Department of Employment, Education and Training (DEET) found that considerable use was being made of evaluation findings. The review, conducted jointly by DEET and DoF, surveyed a large sample of completed evaluations and found that:

- 55% had led to changes in program management;
- 8% had resulted in an improvement in the quality of program outputs;
- 10% had influenced new policy proposals sent by the line Minister to Cabinet for its consideration; but
- there was considerable unevenness within DEET in the quality and use made of evaluation findings — some line areas of DEET had maintained a high standard, while other areas were either consistently poor or were uneven.

Source: Crossfield and Byrne (1994).

surveys of the extent of the influence of evaluation findings on the budget proposals submitted to Cabinet. These have been surveys of DoF officers to seek their judgment concerning the extent of the influence of evaluation. While the survey results provide no more than a broad indication of the extent of influence, they are very revealing.

In the 1990-91 budget, some A\$230 million (then about US\$175 million) of new policy proposals submitted by line Ministers were judged to have been directly or indirectly influenced by the findings of an evaluation. By 1994-95 — the latest year for which estimates are available — this had risen to A\$2300 million. (Measured in dollar terms, the proportion of new policy proposals influenced by evaluation rose from 23% to 77% over that period.) In most cases the influence of evaluation was judged by DoF to be direct.

Some unique features of the 1994-95 budget resulted in these figures being particularly high.⁸ Nevertheless, the results indicate the importance which public servants — in their preparation of the details of new policy proposals — and Ministers have attached to having evaluation findings available. Overall, it has been very important to have had the active support of key Cabinet and other

Ministers in encouraging portfolios to plan and conduct high-quality evaluation.

Evaluation can also have a significant influence on the 'savings options' put forward by DoF or by portfolios for Cabinet consideration in the budget process. (Savings options are areas of government expenditure which could be trimmed or abolished entirely.) In 1994-95 about A\$500 million of savings options — or 65% of the total — were influenced by evaluation findings.⁹

It seems likely that this emphasis on evaluation findings has been encouraged by the nature of the budgetary system in the Australian government. The country has a well-functioning policy-making mechanism which makes transparent the costs of competing policies and encourages debate and consultation among stakeholders within government.¹⁰ In this 'marketplace of ideas' evaluation findings can provide a competitive advantage to those who use them.

One issue which it is important to appreciate is the realistic limits to the influence of evaluation on Ministers' or Cabinet's decision-

making. The evaluation paradigm in an investment project is typically that of cost-benefit analysis: a project is warranted if, but only if, its benefit-cost ratio is greater than one. But program evaluation is a more qualitative science: it can help identify the efficiency or effectiveness of existing, ongoing programs but it can rarely provide an overall conclusion that the activity is worthwhile.

To give an example, if a government decides to allocate a large amount of spending to the unemployed, program evaluation findings can help to map out the probable consequences of alternative types of labor market intervention such as wage subsidies, public sector job creation, or labor market regulation. Program evaluation can be used to identify the cost-effectiveness of each type of intervention. But program evaluation can usually not identify the overall amount of resources which should be allocated. That is a political decision, which is why governments are elected in the first place. The most that program evaluators can realistically and legitimately hope is that their findings are an influential input into government's decision-making and prioritization among competing proposals and programs (see box on next page).¹¹

III. Success Factors and Impediments — What Has/Has Not Worked, and Why

The preceding discussion identified a number of success factors and impediments to success. These are now considered, together with some thoughts on their relative importance.

The Department of Finance (DoF)

DoF has been central to the development of ECD in the Australian government: there have been advantages and some disadvantages in this. Overall, however, DoF's considerable authority with departments and with Cabinet has given it the strength to influence the acceptance of evaluation. This has been probably the single most important factor in the substantial degree of progress with ECD in Australia.

DoF has been an influential, devil's advocate in advising Cabinet about the level of funding which should be allocated to departments for different government programs. As part of this function it provides advice on the new policies proposed by line Ministers, and on possible savings options.

Being the devil's advocate does not endear DoF to other departments, and in fact is an impediment to close cooperation and trust between other departments and DoF. On the other hand, DoF works day-to-day with departments advising on funding and policy issues, participating in reviews and evaluations of their programs, and providing advice on evaluation and other public sector management tools. In evaluation, DoF's evaluation branch with its nine evaluators provided desk officer assistance to departments with advice on methodology, best practice, provision of training courses, publication of evaluation handbooks and guidance material, and support for evaluation networks of practitioners.

The nature of these relationships can vary considerably, with other departments viewing DoF at best as a useful source of advice and treating it with wary respect, and at worst with downright hostility. The former relationship is much more common.

As the evaluation champion, DoF has succeeded in getting evaluation 'on the agenda' in its work with departments. This stands in contrast to the situation which faced the former Office of the

Comptroller-General (OCG) in Canada in the early 1990s. OCG was a stand-alone, specialist body responsible for attempting to influence line ministries to adopt evaluation as a management tool. But OCG was seen as tangential to mainstream government activities, and this undercut its influence. It is interesting to note that OCG's functions have now been relocated as part of the Treasury Board Secretariat, the Canadian equivalent of DoF, to increase its leverage in dealing with line ministries. This relocation was undertaken after a review in 1993 by the Canadian Auditor General of overseas practices, including in particular Australia's.¹²

Another advantage of having DoF responsible for evaluation oversight is that it ensures a direct influence on the line areas of DoF which oversee the line departments. Before the devolutionary reforms of the past fifteen years DoF was heavily involved — some would say 'bogged down' — in the detailed scrutiny of departments' spending activities. The more recent focus on evaluation and other public sector reforms has helped foster a greater focus in these line areas on bottom-line outcomes and value for money; DoF is simply too important a bureaucratic player to allow it to remain with outmoded attitudes and activities. However, achieving this needed cultural change in DoF has taken a number of years, and has involved substantial staff turnover.

The greater focus on value-for-money has also flowed through to the nature and quality of policy advice which DoF provides to Cabinet. That advice has increasingly drawn on available evaluation findings, thus also helping to raise the profile of evaluation with line departments. DoF's involvement in selected evaluations also provides some quality assurance to Cabinet about the evaluation findings on which proposals for new policy might be based.

One example of the evaluative culture which has grown in DoF, and arguably in the then Cabinet itself, was Cabinet's agreement to commission some 60 major reviews of government programs. These reviews had been suggested by DoF in the 1993-94 to 1995-96 budgets, with most focusing on issues of effectiveness and

An Example Of The Use Of Evaluation To Help Government Cut And Reprioritize Its Programs

In the 1996-97 Budget the new government was determined both to reduce and to reprioritize government spending. Particular focus was given to labor market and related programs, which accounted for spending of A\$3,800 million annually (about US \$2,900m).

The Minister for Employment articulated the government's overall policy goal as being to provide assistance to the long-term unemployed and to those at risk of entering long-term unemployment. This focus was adopted both for equity and efficiency objectives — the latter pursued by achieving a better matching of labor supply and demand. At the same time, she wanted to achieve better value for money from labor market programs in the tight budgetary environment.

Australian and international evaluation findings were drawn on heavily to help guide the policy choices made. The Minister highlighted the relative cost-effectiveness of different labor market programs. A key measure of this was estimated by calculating the net cost to government for each *additional* job placement from different programs — as measured by the *increased* probability of an assisted person being in a job 6 months after they had participated in a labor market program. (The baseline was a matched comparison group of individuals who did not participate in a program.)

Evaluation findings showed that the JobStart program, which provides wage subsidies, had a net cost of A\$4,900 per additional job placement, whereas the JobSkills program, which was a direct job creation program, had a net cost of A\$76,600. The Minister noted that “the Government will be ... concentrating its efforts on those programs which have proven most cost-effective in securing real job outcomes”. As a result, the JobStart program was retained while the JobSkills program was substantially scaled back and more tightly targeted to jobseekers who were particularly disadvantaged.

Total savings to the government from its reduction and reprioritization of labor market programs were about A\$1,500 million over two years.

Cabinet also commissioned a series of major evaluations of its new labor market programs and of the new arrangements for full competition between public and private employment service providers.

Source: Senator Vanstone (1996); DEETYA (1996, 1997); Commonwealth of Australia (1996).

appropriateness; many of these reviews surveyed and summarized existing evaluative information, rather than conducted in-depth evaluations themselves.

The reviews related to aspects of programs which collectively involved about A\$60 billion in annual expenditure. These reviews were designed as an urgent response to emerging budget pressures, and might best be viewed as complementary to the regular cycles of evaluation as reflected in portfolio evaluation plans. Such Cabinet-endorsed reviews can be a useful vehicle for a DoF to use if

line departments strongly resist DoF suggestions about evaluation priorities.

One benefit to line departments from DoF involvement in individual evaluations is that DoF can draw on evaluation skills and experience, spanning the whole breadth of government activities. Most DoF officers are usually not specialists in technical evaluation issues — they are expenditure and financial policy specialists. Perhaps their greatest potential value-added is the objective approach which they can bring to bear to an evaluation — DoF

officers are comfortable in asking difficult questions about a program's performance, and in steering an evaluation towards these issues. This independent and questioning approach has provided a useful counterpoint to the in-depth knowledge (but often partisan approach) of program areas within line departments.

Line departments

Most of the day-to-day work of line departments relates to ongoing program management. One of the three objectives of the evaluation strategy is to encourage program managers to use evaluation to improve their programs' performance. This has proved surprisingly difficult at times.

To those who understand the potential contribution of evaluation its utility seems almost self-evident. Evaluators often plead along the lines 'How can you program managers improve your programs and better meet the needs of your clients unless you carefully evaluate program performance?' Unfortunately, appeals to the professionalism of senior executives are not terribly effective. Most managers understand the potential benefits of evaluation, but do not place it at the top of their day-to-day priorities, particularly in a climate of tightly constrained and diminishing resources.¹³

Experience within line departments in Australia indicates that a highly supportive culture is necessary if major evaluations are to be planned, resources allocated to properly manage and undertake them, and the findings implemented. The commitment of departmental secretaries (the CEO) to achieving improvement in program performance is paramount in fostering such a results-based management culture.

Over at least the past decade, the tenure of secretaries has often been brief. This turnover has meant that some departments have had a series of secretaries who have placed varying priority on evaluation and the departmental effort devoted to it. While an evaluative culture can be slow to build up, it can and has been reduced much more rapidly.

Reprioritization of labor market programs provides one, high-profile example of the potential benefits of evaluation to Ministers and their departments. More generally, there has been an emphasis in recent years on 'portfolio budgeting'. This includes the setting by

Cabinet of portfolio spending targets at the start of each annual budget round. In the tight budgetary environments that have predominated, the nature of these targets usually imply that if Ministers wish to propose any new policies then these must be funded from within the portfolio's spending envelope. Evaluation has been one tool to assist Ministers and their secretaries in the design of new policies and in the prioritization among existing policies (subject, of course, to Cabinet's endorsement).

A portfolio budgeting approach helps to ensure that the focus of Ministers and their departmental secretaries is on value-for-money issues, as well as management-oriented efficiency issues. Thus portfolio budgeting is a key part of the devolutionary public sector reforms in Australia.

The requirement for portfolio evaluation plans has necessitated that departments set up a bureaucratic infrastructure to prepare them (which may be a flaw when the government is determined to cut red tape and the bureaucracy, as at present). In about three-quarters of departments this has involved the creation of a committee, usually chaired by a deputy secretary of the department, to meet regularly, canvass candidate programs for future evaluation, and monitor the progress of evaluations already underway.¹⁴ This work itself generates bureaucratic momentum. Most departments involve their Minister and their Minister's office by seeking their comments on (and clearance of) the draft evaluation plans.

It is difficult to speculate with any confidence how the evaluation 'scene' in Australia would have looked in the absence of a powerful champion such as DoF. Some of the larger departments, such as the Departments of Employment and Health, would no doubt have had a substantive evaluation effort in any event — the evaluation emphasis in the Department of Employment pre-dated the government's formal evaluation strategy. However, informal discussions with senior executives in those departments have emphasized the catalytic influence of DoF even in their departments. Executives responsible for the central evaluation areas in line departments have generally found DoF a natural ally in helping to persuade more traditional administrators in their departments to adopt evaluation as a valued management tool.

Most departments have chosen to rely on program managers and their staff for the actual conduct of evaluations. This devolutionary

approach has helped to ‘mainstream’ evaluation as a core activity of each line area, and has ensured that evaluations draw heavily on the program knowledge and experience of those who actually manage the program. It has also led to a greater appreciation of the complementarity — and sometimes the substitutability — between in-depth program evaluation and the more frequent monitoring of performance via the collection of ongoing performance information.

The devolutionary approach has also secured ‘ownership’ by program managers of the evaluation findings. These are important advantages, and provide a strong contrast with externally-conducted evaluations, reviews or performance audits where lack of program knowledge and commitment to implement the findings has often significantly undermined the impact of findings.

But there have also been disadvantages to this devolved approach. One has been a lack of evaluation skills in many program areas and inexperience in conducting evaluations, (as suggested by the recent survey by the ANAO of a sample of evaluation reports).

Basic training in evaluation skills is widely available in the Australian government — provided by DoF in particular¹⁵ — as is DoF and departments’ own guidance material such as evaluation handbooks. There is a substantial community of evaluation consultants in Canberra, including numerous academics with either subject area knowledge (such as health issues) or with specialist research and analysis skills. However, the recent ANAO study showed that 20% of departments are concerned about the lack of available training in advanced evaluation techniques.

Some departments have addressed the need for more advanced skills and experience by setting up a central evaluation unit to provide advice on methodology and to participate in evaluation steering committees. The Department of Health has pursued evaluation quality assurance in a devolved environment via ensuring that adequate skills and resources are available to program managers together with structural arrangements in place, such as technical panels and steering committees.¹⁶ That department, like others, puts a lot of effort into training its staff to enhance their analytical and research skills.

Another disadvantage to the devolved approach is that program staff are often too close to their program to view it objectively and

to ask the hard, fundamental questions concerning its performance and the need for the program to continue. External participation by a central evaluation unit or by peers from other programs in working groups or steering committees has been one way to address this. External participation has often included DoF for major evaluations, and this has been another means of fostering objectivity and rigor.

In only one agency, the Aboriginal and Torres Strait Islanders Commission (ATSIC), is there a separate evaluation body — the Office of Evaluation and Audit (OEA) — which has statutory independence. The independence of the OEA helps to answer claims of ethical difficulties and corruption in the administration of some ATSIC programs. A body such as OEA can be effective in ensuring that accountability objectives are met. The impact of its evaluations and reviews may have been reduced, however, by perceptions that OEA lacks fluency in program understanding and has not secured ownership by ATSIC program managers.

Australian National Audit Office (ANAO)

The ANAO has been a central presence since the inception of the formal evaluation strategy. In endorsing the strategy, Cabinet agreed with the proposition that ‘it is expected that (the ANAO) would contribute to the proposed evaluation strategy through audits of evaluation processes within departments and agencies, including their follow-up on evaluation findings’.

Since 1990 the ANAO has pursued this ‘sheepdog’ task vigorously, both with respect to line departments and to DoF. (It is notable that the Canadian Auditor General has worked in the same way.)¹⁷ The ANAO has conducted seven performance audits into the evaluation and performance information practices of a number of departments during that period, as well as on the overall, government-wide progress with the evaluation strategy. It has also published two “good practice” guides, one of which was prepared jointly with DoF.

In addition to these reviews of evaluation and performance information, the ANAO has placed increasing emphasis on the conduct of performance audits into the economy, efficiency and effectiveness of programs. The ANAO completes about 40 performance audits annually, and these now account for about one half of the Office’s overall activity. The ANAO takes care to ensure that its

performance audit activities — which can be regarded as a form of evaluation¹⁸ — do not overlap or duplicate those of departments, and departments and DoF avoid duplicating ANAO's audit activities when planning their own evaluation priorities.¹⁹

The impact of the ANAO's activities has been felt in several ways. First, it has focused attention on evaluation as a legitimate and important area for senior management attention in departments.

A different impact was felt in the earlier years, when the ANAO pursued its performance audits into evaluation (and into program administration more generally) with a 'gotcha', fault-finding zeal. The value-added of such an approach is highly doubtful as it strongly discouraged the 'victim' departments from ownership of the audit findings. The resistance of these departments to accepting the ANAO findings was often evident in their formal, published responses to ANAO reports.

A 'gotcha' approach may have satisfied a narrow interpretation of the accountability function of an audit office, particularly in its reporting to parliament, but it undermined the potential value-added contribution which a considered performance audit could provide to a line department's future management of a program.

In more recent years, with a new Auditor-General and a different audit philosophy in the Office, there has been a much stronger emphasis on finding ways to help departments improve their performance. A high priority has also been attached to the identification and sharing of good practices, and the ANAO has been active in disseminating these among departments.

Other factors

Having a government-wide evaluation effort, involving all departments, has proved helpful in developing a general climate of expectation that evaluation will be conducted and used, and in developing an evaluation community, especially in Canberra. It has also helped to develop a labor market for evaluation skills, including advanced data analysis skills. The labor market includes the growing number of staff with experience in evaluation units, in the various economic research bureaus, and in the national statistical agency.

One expression of this community has been the monthly meetings of the Canberra Evaluation Forum. The meetings have been organized by

a steering group of departments, with DoF support, and each meeting involves several speakers and discussants of topical evaluation issues. About 100 participants attend each month.

The sharing of insights and good practices through these meetings has strongly encouraged networking. There have been several special-interest conferences and seminars on particular evaluation issues organized by DoF and others, on issues such as the evaluation of policy advice and evaluation/audit links.

A feature of the Australian scene has been the frequent availability of commercially-organized, for-profit conferences on evaluation and other performance measurement issues, and on public sector reform issues more broadly. Various departments including DoF work collaboratively with conference organizers to identify topical issues and provide speakers. The conferences allow an opportunity for federal public servants to be exposed to evaluation issues in state governments, local government and the private sector, and academia. However, the contribution of academia to ECD in the federal government has been more limited than might have been expected.

The role of parliament has not lived up to the ambitious aims of the designers of the evaluation strategy, who viewed accountability to parliament as one of the foundations of the public sector reforms, including the evaluation strategy. In practice, parliament has generally possessed neither the infrastructure resources nor the perspective to focus on the insights into program performance which evaluation findings can offer. While parliament exercises general oversight, including oversight of annual appropriations, it has provided little scrutiny of strategic issues of performance, preferring instead to focus on administrative errors which might embarrass the government.

But there have been some notable exceptions to the narrow focus and interests of parliament, and these have involved parliamentary committees inquiring into issues such as the Financial Management Improvement Program and service quality.²⁰ These have been useful in emphasizing performance issues, such as the impact of government programs on their ultimate clients, to departments and to public servants more generally.

Finally, a number of success factors are often taken for granted in Australia, which become evident when making cross-national

comparisons, particularly with developing countries which do not possess a well-developed public service infrastructure:

- strong institutional and human capacity in the public sector;
- well-developed management capacity;
- public service managers with a reputation for integrity, honesty and impartial advice;
- a well-developed budget management system, and accounting standards and systems;
- a tradition of transparency and accountability in the conduct of government business; and
- a credible and legitimate political executive.

IV. Current Developments And Prospects

The government, elected in early 1996, expressed considerable unhappiness with the federal public service, and considers it to be rule-bound and caught up in red tape. The government has a strong ideological preference for the private sector, and appears to regard it as being inherently more efficient than the public sector. The government has noted with dismay a major evaluation which has shown public service administrative efficiency to be lagging considerably that of private sector best practice, particularly in personnel practices, and this comparison has strengthened its resolve to seek a smaller, more efficient public sector.²¹ The government has, therefore, embarked on a wave of major public sector management reform.

The new Cabinet has directed that Ministers and their departments review the nature and extent of continuing need for existing government programs; and it has expressed the expectation that competitive tendering and contracting processes be applied by departments to their programs wherever possible — this could result in considerable outsourcing of the delivery of government programs in coming years. These types of review require a close scrutiny of performance to be successful. In particular, the application of competitive tendering and contracting necessitates a clear understanding of program objectives, followed by ex-ante assessments of the performance of alternative tenders (in-house and external). Once contracts have been let, there is a need for ongoing performance scrutiny, and at the completion of contracts a review of past performance; this information then feeds back into decisions about the next round of contracts to be put out to tender. Evaluation of performance is central to all these activities.

The government's initiatives are part of a strong push towards commercialization and the private sector delivery of public services, and have already resulted in a significant reduction in the number of public servants, with further larger reductions in prospect. Under some authoritative scenarios, the size of the public service in a decade could be only a fraction of its recent levels.

The government is taking steps to achieve sharper accountability for public service managers, together with fewer centralized controls and fewer formal requirements, partly via a strongly devolutionary approach. Departmental CEOs, who are on employ-

ment contracts, will be required to perform to high standards. These expectations have recently been made explicit in new legislation on financial management and accountability, and will increase pressure on CEOs to ensure high standards of corporate governance.²² This may also create both the scope and a requirement for the ANAO to take a greater quality assurance role in performance measurement and reporting than in the past.

Service delivery agencies will be required to set explicit customer service standards, with actual performance being reported publicly, including to parliament. The focus on performance will be further enhanced by the government's decision to adopt accrual accounting — this will facilitate scrutiny and benchmark comparisons of departmental costs and performance. Changes to output/outcomes reporting are also in prospect, and these will seek to marry the output specification and focus of governments such as those of New Zealand and several of the Australian states and territories, with the outcomes and performance focus of the federal Australian government. This development would invite closer scrutiny of departments' planned and actual performance.

Collectively, the latest wave of reforms are likely to result in very considerable changes in performance management, accountability and reporting. For these reforms to be successful, however, there will need to be a high level of scrutiny of departmental and CEO performance to further strengthen management incentives.

This environment helps explain and put in context Cabinet's recent agreement to the replacement of the formal requirements of the evaluation strategy by a principles-based approach. This emphasizes the uses of evaluation and other performance information for performance management purposes, including links with corporate and business planning and the other reform initiatives now underway. Thus the approach in one sense reaffirms the merits of the Financial Management Improvement Program, and of program budgeting. The new approach continues to emphasize the advantages in planning, conducting, reporting and using evaluation findings, the main difference with the previous evaluation strategy now being the absence of formal requirements.²³

How should the new approach be viewed? If at least part of the

success of the requirement-based approach was because it mandated evaluation activity, then there will be some risk with a new approach which intentionally excludes such formal requirements.

But a counter-argument is that the focus of concern should be with outcomes, not with processes to achieve them. If the public sector management framework provides sufficient incentives to achieve a strong focus on performance and outcomes, then this provides support to devolutionary approaches which provide management with the autonomy to achieve this performance in any manner it chooses. In a government where performance measurement has been strengthened, and there is greater accountability for results, there is scope to provide departments with greater autonomy and flexibility.²⁴

The new approach to evaluation accepts performance measurement as an integral part of performance management²⁵ —

reflecting a philosophy that if the environment of public sector governance is strongly conducive to evaluation being conducted and used, then that will happen. Thus the emerging public sector environment would be expected to be even more encouraging of evaluation than in previous years. But this expectation might mirror a similar but erroneous one in the early 1980s, when it was assumed that if the structural framework of public service management was 'correct', then an evaluative culture would almost automatically follow.

The impact of the new reforms on the culture and management of the public service will partly depend on progress already achieved since the early 1980s. To the extent that an evaluation culture — including management commitment to review and learn from past performance, and an evaluation infrastructure — has already been achieved, this will enhance the speed and extent of impact of the new reforms.

V. Conclusions From The Australian Experience

The requirements-based, formal evaluation strategy in Australia constituted a model of central force-feeding to ensure the planning, conduct, quality and use of evaluation. It reflected the belief that managers would not do this if left to their own devices. Formal requirements and a whole-of-government approach helped to kick-start the process and achieve significant momentum, but it is worth bearing in mind that ‘winning hearts and minds’ cannot be mandated. It is not formal rules and requirements that determine the extent of the conduct and use of evaluation findings — it is the commitment of individuals and their organizations, and the nature of their understanding and motivation.

The recent move to a principles-based approach reflects the evolution of governance arrangements and the particular circumstances of the Australian scene. This evolution represents a migration from tight to loose controls over departments.

The Australian experience provides a wealth of lessons. However, although it is possible to identify features of that system which have contributed to the substantial success achieved so far, this does not necessarily mean that these success factors are *preconditions* for success.

Some of the key success factors have been:

- macroeconomic pressures which have led to tight budgets and a priority on finding ways of achieving better value-for-money;
- a powerful department (DoF) willing to champion evaluation, react to changing circumstances and identify new opportunities for influence and development;
- the sustained commitment over a decade of the government, and especially of its main champion (DoF), to the evaluation strategy;
- having a second central agency (the ANAO) willing to prompt and prod departments to focus on evaluation and performance management more broadly;
- the creation of an explicit evaluation strategy with formal evaluation requirements;
- a whole-of-government strategy to help achieve and maintain momentum in evaluation capacity development;
- a budget agency (DoF) able to link evaluation into both the budget process and into public sector management reforms;
- a budget system which makes transparent the costs, and the pros and cons of competing policies;
- the implementation of related public sector management reforms, particularly portfolio budgeting, which provide substantial autonomy to line managers and which emphasize bottom-line results and outcomes ³/₄ these have provided powerful incentives to managers;
- the support of Cabinet and a number of key Ministers, and the emphasis they have placed on having evaluation findings available to assist their decision-making;
- the priority given to evaluation in several large and important line departments, which has helped to highlight and legitimize it; and
- the devolutionary approach to evaluation within line departments, which has helped to mainstream evaluation as one of their core activities, together with internal quality assurance processes.

VI. Some Implications For The World Bank

A challenge for the Bank is to foster ECD in developing countries. So what lessons can be drawn from the Bank's own experience in fostering ECD, and to what extent are these similar to the Australian experience?

The report of the Bank's 1994 Task Force on ECD diagnosed a number of problems in developing and embedding evaluation capacity in developing countries.²⁶ These included:

- a lack of genuine demand and ownership by politicians and officials — demand was judged by the Bank to be the key precondition;
- lack of a culture of accountability — often reflecting problems of ethics and corruption;
- lack of necessary evaluation, accounting and auditing skills. Overcoming this often requires broader institutional development and capacity-building;
- poor quality of financial and other performance information, and of the accounting and auditing standards and systems required to provide and make use of such information;
- lack of evaluation feedback mechanisms into decision-making processes; and
- the need for ECD efforts to have a minimum critical mass if they are to succeed.

Lack of real demand has continually been identified as the crucial deficiency. But what does this mean? It would be unrealistic to expect wholehearted support across an entire public service in any country — developing or developed — or from all government ministers.

One strategy could be to foster evaluation in only one or two departments, and to hope that the demonstration effect would cause other departments to progressively adopt performance measurement approaches too. This 'enclave' approach could work — good practice examples are invaluable in demonstrating the potential of evaluation. But efforts could be set back considerably whenever there are changes in senior departmental management, or when policy or funding crises cause a diversion of focus. ECD efforts in some developing countries have stalled for exactly these reasons.²⁷ Experience in Australian departments shows that

performance measurement initiatives, and other public sector reform initiatives, can be the first to be postponed — sometimes indefinitely — when external pressures occur.

In contrast, a government-wide approach offers the potential to generate sufficient momentum to sustain progress in all departments. Even departments which suffer some external setbacks can be induced to keep up with their peers if there is sufficient government-wide pressure and momentum.

A government-wide approach requires at least one strong, lead department or agency $\frac{3}{4}$ perhaps ideally two. Central agencies, such as Ministries of Finance or Planning, or a National Audit Office, are prime candidates. Issues to consider here when examining possible champion agencies include the depth and sustainability of their commitment, and their ability to prod and support other agencies effectively. An ideal champion could influence both public expenditure management and line management within all other agencies.

Cabinet or government endorsement is a powerful lever, but this should be viewed in context. A government would be unlikely to view evaluation as anything more than a useful tool; the Australian experience is that government ministers often regard evaluation as bureaucratic business — i.e., something for officials to focus on. However, if they are advised by their senior officials that evaluations should be commissioned and that this will assist policy formulation and decision-making, then Ministers have been happy to endorse them.

While formal Cabinet endorsement of an evaluation strategy has been an important lever, it is necessary to have a strong lead department to champion the strategy with other departments. Unless this happens evaluation will only be paid lip service.

Another important environmental factor is a sense of urgency such as a budgetary crisis. This can help to persuade officials and their Ministers of the need for a systematic approach to performance measurement.

In the absence of such an environment it should be possible to

pursue ECD as part of broader public sector management/governance efforts. The Bank has argued that ECD should be viewed as an integral part of these broader efforts, but evaluation is often seen as a stand-alone activity. There is often insufficient understanding even among PSM reformers that evaluation is an invaluable and necessary support to policy analysis, budgetary resource allocation, and to program and organizational management discrete projects and of ongoing programs. A stronger version of this proposition is that high levels of economy, efficiency and effectiveness are unattainable unless there are sound and integrated systems of performance measurement and management.

There are potential synergies and commonalities between Bank work on ECD and its support for broad governance efforts, including areas such as civil service reform, financial reporting and auditing, and anti-corruption efforts. And if developing country governments place increasing emphasis on outsourcing the delivery of government activities this will provide an additional opportunity for them to move to a more evaluative culture — however, this will also require that they possess sound assessment and contract management skills.

Given that it takes at least a decade to develop a national evaluation system and embed it in a government in a sustainable manner, there are implications for Bank and other development agency support for ECD. Clearly, it is necessary for the Bank to move away from a short-term project focus in the technical assistance it provides, towards one which gives strong and enduring support over the long-term.

The position of the World Bank is analogous in some ways to that of the Australian DoF — the Bank as a catalyst and advocate with developing country governments. The lesson from Australia is that a focused approach, with substantive and sustained momentum, is necessary to overcome internal lack of support and to exploit external opportunities whenever these arise or can be created.

Bibliography

- ANAO (Australian National Audit Office). 1991. *Implementation of Program Evaluation — Stage 1*. Efficiency Audit Report No. 23, 1990-91. Canberra: Australian Government Publishing Service (AGPS).
- _____. 1991. *Evaluation in Preparation of the Budget*. Efficiency Audit Report No. 13, 1991-92. Canberra: AGPS.
- _____. 1992a. *Program Evaluation in the Departments of Social Security and Primary Industries and Energy*. Efficiency Audit Report No. 26, 1991-92. Canberra: AGPS.
- _____. 1992b. *Auditing Program Evaluation — ANAO Performance Auditing Guide*. Canberra: AGPS.
- _____. 1992c. *Department of the Treasury — Procedures for Managing the Economic Policy Program*. Efficiency Audit Report No. 36, 1991-92. Canberra: AGPS.
- _____. 1993. *Program Evaluation — Strategies, Practices and Impacts — Industry, Technology and Regional Development Portfolio*. Efficiency Audit Report No. 35, 1992-93. Canberra: AGPS.
- _____. 1996. *Performance Information — Department of Employment, Education, Training and Youth Affairs*. Performance Audit Report No. 25, 1995-96. Canberra: AGPS.
- _____. 1997a. *Applying Principles and Practice of Corporate Governance in Budget Funded Agencies*. Canberra: AGPS.
- _____. 1997b. *Program Evaluation in the Australian Public Service*. Performance Audit Report No. 3, 1997-98. Canberra: AGPS.
- ANAO/DoF (Department of Finance). 1996. *Performance Information Principles*. Canberra: ANAO.
- Auditor General of Canada. 1993. *Report of the Auditor General of Canada to the House of Commons. Chapter 8: Program Evaluation in the Federal Government; Chapter 9: Operation of Program Evaluation Units; Chapter 10: The Program Evaluation System — Making it Work*. Ottawa.
- _____. 1996. *Report of the Auditor General of Canada to the House of Commons. Chapter 3: Evaluation in the Federal Government*. Ottawa.
- _____. 1997. *Report of the Auditor General of Canada to the House of Commons. Chapter 5: Reporting Performance in the Expenditure Management System; Chapter 11: Moving Toward Managing for Results*.
- Barrett, Pat, Auditor-General of Australia. 1996a. “Some Thoughts About the Roles, Responsibilities and Future Scope of Auditors-General.” Address at the Australian Society of Practising Accountants Annual Research Lecture, Canberra, November 14.
- _____. 1996b. “Performance Standards and Evaluation.” Address at the Institute of Public Administration of Australia National Conference, Reshaping the Old: Charting the New — Public Management in the 1990s, Melbourne, November 20-22.
- Commonwealth of Australia. 1996. *1996-97 Budget Statement 3: Overview*. Canberra: AGPS.
- Crossfield, Len and Anne Byrne. 1994. *Review of the Evaluation Function in DEET*. Canberra: Department of Employment, Education and Training, and Department of Finance.
- DEETYA (Department of Employment, Education, Training and Youth Affairs). 1996. *Budget Initiatives*. Canberra: DEETYA.
- _____. 1997. *Enhancing the Effectiveness of Active Labour Market Policies*. Australian presentation to the meeting of OECD Labour Ministers. Canberra: DEETYA.
- Dixon, Geoff. 1993. “Managing Budget Outlays 1983-84 to 1992-93.” In Brian Galligan, ed., *Federalism and the Economy: International, National and State Issues*. Canberra: Federalism Research Centre, Australian National University.
- DoF (Department of Finance). 1988. *FMIP Report*. Canberra: AGPS.
- _____. 1991. *Handbook of Cost-Benefit Analysis*. Canberra: DoF.
- _____. 1993. “The Cost of Evaluations: the Findings of a Pilot Study.” DoF Working Paper, Canberra.
- _____. 1994a. *Doing Evaluations — a Practical Guide*. Canberra: DoF.
- _____. 1994b. *The Use of Evaluation in the 1994-95 Budget*. Finance Discussion Paper, DoF, Canberra.
- _____. 1995. “Reasons Why Evaluation Should Be Done and Why Finance Should Be Involved.” DoF Working Paper, Canberra.
- _____. 1996. *Performance Information Review: Overview of the First Year of the Review*. Canberra: DoF.
-

-
- DoF and Public Service Board. 1986. *Evaluating Government Programs — A Handbook*. Canberra: AGPS.
- Duckett, Mary. 1995. *Performance Reporting in Commonwealth Annual Reports*. Report for the Department of Finance. Canberra: DoF.
- Funnell, Sue. 1993. *Effective Reporting in Program Performance Statements*. Study for the Department of Finance. Canberra: DoF.
- Keating, Mike, and Malcolm Holmes. 1990. "Australia's Budgeting and Financial Management Reforms." *Governance: an International Journal of Policy and Administration* 3(2): 168-185.
- Mackay, Keith. 1994. "The Australian Government's Evaluation Strategy: a Perspective From the Center." *Canadian Journal of Program Evaluation* 9(2): 15-30.
- MAB/MIAC (Management Advisory Board/Management Improvement Advisory Committee). 1995. *Achieving Cost Effective Personnel Services*. MAB/MIAC Report No. 18. Canberra: AGPS.
- Parliament of Australia, House of Representatives Standing Committee on Finance and Public Administration. 1990. *Not Dollars Alone — Review of the Financial Management Improvement Program*. Parliament of Australia, Canberra.
- Parliament of Australia, Senate Finance and Public Administration References Committee. 1995. *Service Delivery: Report from the Committee on Service Delivery by the Australian Public Service*. Parliament of Australia, Canberra.
- President of the Treasury Board. 1995. *Strengthening Government Review — annual report to the Parliament*. Ottawa: Treasury Board Secretariat.
- Reith, Peter, Minister for Industrial Relations and Minister Assisting the Prime Minister for the Public Service. 1996. *Towards a Best Practice Australian Public Service*. Discussion Paper. Canberra: AGPS.
- Schick, Allen. 1996. *The Spirit of Reform: managing the New Zealand state sector in a time of change*. Report prepared for the New Zealand State Services Commission and the Treasury. Wellington: State Services Commission.
- _____. 1998. "Why Most Developing Countries Should Not Try New Zealand Reforms." *The World Bank Research Observer*. 13(1): 123-131.
- Sedgwick, Steve. 1996. "Setting Ourselves for the Future." Address to DoF staff. Canberra: DoF.
- Senator Amanda Vanstone, Minister for Employment, Education, Training and Youth Affairs. 1996. *Reforming Employment Assistance — helping Australians into real jobs*. Canberra: AGPS.
- Task Force on Management Improvement. 1992. *The Australian Public Service Reformed: an Evaluation of a Decade of Management Reform*. MAB/MIAC, Canberra.
- Uhr, John and Keith Mackay, eds. 1996. *Evaluating Policy Advice: Learning From Commonwealth Experience*. Canberra: Federalism Research Centre (Australian National University) and DoF.
- World Bank. 1994. *Evaluation Capacity Development — Report of the Task Force*. World Bank, Washington, D.C.
- _____. 1997. *The State in a Changing World: World Development Report 1997*. New York, N.Y.: Oxford University Press.

Endnotes

¹ This paper was presented to a World Bank seminar on *Public Sector Performance - the Critical Role of Evaluation*, in April 1998.

* I am happy to acknowledge the invaluable comments and suggestions of Ray Rist, Ladipo Adamolekun, Len Crossfield and Martin Mullane.

² The three dimensions on which program evaluation can focus are (i) the efficiency of a program's operations (minimizing costs for a given level of output), (ii) its effectiveness in achieving its objectives, and (iii) whether the program's objectives remain consistent with the government's policy priorities, (the appropriateness of the program).

³ There was only modest success with the requirement that Ministers' new policy proposals include an evaluation plan of action that would be undertaken if the proposal was accepted. Feedback from portfolios indicated that this requirement was onerous for portfolio managers during the busy budget period. Only about 30% of proposals broadly met this requirement in the 1993-94 budget, for example, although an additional 50% of proposals included a clear undertaking to evaluate the proposal if accepted (DoF 1994b). These percentages were only achieved after considerable prodding by line areas within DoF. In recent years the extent of such prodding (and of departments' willingness to provide such plans in their budget documentation) has fallen off considerably.

⁴ ANAO 1997b.

⁵ Ibid.

⁶ The ANAO (ibid) recently surveyed 20 departments and agencies and found that 11 had reviewed their evaluation activities within the past four years.

⁷ Ibid.

⁸ A major review of labor market policies resulted in a particularly heavy emphasis on evaluation findings in that budget (see DoF 1994b).

⁹ This compares with A\$1060 million in the preceding year (reflecting a composition effect because of small numbers of large savings options).

¹⁰ The Australian budgetary system is discussed in World Bank 1997, p. 82.

¹¹ In the absence of evaluation findings, decisions will probably be influenced more by ex-ante analysis or anecdotal information and case study examples.

¹² Auditor General of Canada 1993.

¹³ A 1992 evaluation of public sector management reform in Australia concluded that "there is widespread acceptance of the importance of evaluation". But it went on to note that "the bulk of (senior executive managers) state that it is using evaluation information only sometimes or infrequently during the conduct of their job. This suggests that information generated by evaluations is not yet a key element in program management". (Task Force on Management Improvement 1992, pp. 378 and 379.)

¹⁴ See ANAO 1997b.

¹⁵ DoF has provided introductory evaluation training to over 3000 public servants since 1991.

¹⁶ The Department of Health encourages quality evaluations through: selection of good quality officers to manage the evaluation; involvement of internal and external stakeholders; ensuring that technical advisory panels are available to help assess the work of consultants; having steering groups available to help manage consultants; and ensuring that sufficient resources are available for the evaluation.

¹⁷ See, for example, Auditor General of Canada 1996 and 1997.

¹⁸ A broad definition of evaluation is used in Australia. It includes program evaluation, project evaluation (principally cost-benefit analysis), efficiency and performance audits, and formal policy reviews.

¹⁹ Auditor-General 1996; ANAO 1997b.

²⁰ Parliament of Australia 1990, 1995.

²¹ Reith 1996; MAB/MIAC 1996. The government has favorably acknowledged the existence of "a system of performance management and program budgeting based upon an explicit evaluation of outcomes". (Reith 1996, p.ix)

²² The new legislation was passed by parliament in October 1997. See also ANAO 1997a.

²³ And yet even some of the formal requirements will remain, but in a different guise: departments will continue to be issued guidelines for reporting to parliament about their annual appropriations, and these will now include the need for summary statements of evaluation intentions (i.e. plans). In addition, guidelines for preparation of departments' annual reports will note the need to report on past performance, including results as shown by completed evaluations and other performance information.

²⁴ This approach is advocated by the World Bank in its recent World Development Report (World Bank 1997).

²⁵ In contrast, it could be argued that the earlier requirements-based approach had been added on to the then-existing suite of reform initiatives in the 1980s almost as an afterthought. If that is a fair interpretation, it was a very effective afterthought.

²⁶ World Bank 1994.

²⁷ Ibid.

Other Forthcoming Papers in this Series

Keith Mackay, *Lessons From National Experience*.

Stephen Brushett, *Evaluation Capacity Development in Zimbabwe*.

Alain Barberie, *Indonesia's National Evaluation System..*

Other Recommended Reading

1. The World Bank, *Colombia: Paving the Way for a Results-Oriented Public Sector*, WB Report no. 15300-CO, January 28, 1997.
2. The World Bank, *Report of the Evaluation Capacity Task Force*, June 1994.
3. Eduardo Wiesner D., *From Macroeconomic Correction to Public Sector Reform: the Critical Role of Evaluation*, WB Discussion Paper no. 214, 1993.
4. Derek Poate, *Measuring & Managing Results: Lessons for Development Cooperation*, SIDA/UNDP, 1997.
5. Robert Picciotto and Eduardo Wiesner D. (eds.), *Evaluation and Development*, Transaction Publishers, New Jersey, 1998.
6. R. Pablo Guerrero O., 'Evaluation capacity development in developing countries: applying the lessons from experience', in R. Boyle and D. Lemaire (eds.), *Building Evaluation Capacity*, Transaction Publishers, New Jersey, 1998.

Operations Evaluations Department Publications

The Operations Evaluations Department (OED), an independent evaluation unit reporting to the World Bank's executive directors, rates the development impact and performance of all the Bank's completed lending operations. Results and recommendations are reported to the executive directors and fed back into the design and implementation of new policies and projects. In addition to the individual operations and country assistance programs, OED evaluates the Bank's policies and processes.

Operations evaluation studies and World Bank discussion papers can be obtained through the World Bank bookstore. All other documents are available from the World Bank Public Information Center (PIC).

Summaries of studies and the full text of the *Précis and Lessons & Practices* can be read on the Internet at <http://www.worldbank.org/html/oed/index.htm>

How To Order OED Publications

Documents listed with a stock number and price code may be obtained through the World Bank's mail order service or from its InfoShop in downtown Washington, DC. For information on all other documents contact the World Bank InfoShop.

Ordering World Bank Publications

Customers in the United States and in territories not served by any of the Bank's publication distributors may send publication orders to:

The World Bank
P.O. Box 960
Herndon, VA 20172-0960
Fax: (703) 661-1501
Telephone: (703) 661-1580

The address for the World Bank publication database on the Internet is: <http://www.worldbank.org> (click on publications).

Internet: <http://www.worldbank.org/>. From the World Bank homepage, select *publications*.

E-mail: pic@worldbank.org

Fax number: (202) 522-1500

Telephone number: (202) 458-5454

The World Bank InfoShop serves walk-in customers only. The InfoShop is located at:

701 18th Street, NW
Washington, DC 20433, USA

All other customers must place their orders through their local distributors.

Ordering via E-Mail

If you have an established account with the World Bank, you may transmit your order via electronic mail on the Internet to: **books@worldbank.org**. Please include your account number, billing and shipping addresses, the title and order number, quantity, and unit price for each item.

The OED Working Paper Series is produced by OEDPK (Partnerships and Knowledge). Copies or inquiries about these or other publications in our series can be obtained by calling 202-473-5365 or email: ecampbell-page@worldbank.org.

If you would like to receive a complimentary copy of the World Bank's forthcoming publication on *Public Sector Performance - the Critical Role of Evaluation*, 1998, please contact Ms Elaine Wylie via email (Bwylie@worldbank.org) or phone 202-473-1718.

OED Study Series

- India: The Dairy Revolution
- The World Bank's Experience with Post-conflict Reconstruction (1998)
- Financial Sector Reform: A Review of World Bank Assistance (1998)
- Bangladesh Country Assistance Review (1998)
- Agricultural Extension and Research: Achievements and Problems in National Systems (1997)
- Fiscal Management in Adjustment Lending (1997)
- Reforming Agriculture: The World Bank Goes to Market (1997)
- Paddy Irrigation and Water Management in Southeast Asia (1997)
- 1995 Evaluation Results (1997)
- Zambia Country Assistance Review: Turning an Economy Around (1997)
- The Aga Khan Rural Support Program: A Third Evaluation (1996)
- Lending for Electric Power in Sub-Saharan Africa (1996)
- Industrial Restructuring: World Bank Experience, Future Challenges (1996)
- Social Dimensions of Adjustment: World Bank Experience, 1980-93 (1996)
- 1994 Evaluation Results (1996)
- Ghana Country Assistance Review: A Study in Development Effectiveness (1995)
- Evaluation and Development: Proceedings of the 1994 World Bank Conference (1995)
- Developing Industrial Technology: Lessons for Policy and Practice (1995)
- The World Bank and Irrigation (1995)
- 1993 Evaluation Results (1995)
- Structural and Sectoral Adjustment: World Bank Experience, 1980-92 (1995)
- Gender Issues in World Bank Lending (1995)
- The World Bank's Role in Human Resource Development in Sub-Saharan Africa: Education, Training, and Technical Assistance (1994)
- 1992 Evaluation Results (1994)
- New Lessons from Old Projects: The Workings of Rural Development in Northeast Brazil (1993; contains summaries in French, Portuguese and Spanish)
- World Bank Approaches to the Environment in Brazil (1993; contains summaries in French, Portuguese, and Spanish)
- Trade Policy Reforms under Adjustment Programs (1992)
- World Bank Support for Industrialization in Korea, India, and Indonesia (1992)
- Population and the World Bank: Implications from Eight Case Studies (1992)
- The Aga Khan Rural Support Program in Pakistan: Second Interim Evaluation (1990)