



8

Australia

Australia elected a reformist government in 1983. That government faced a very difficult macroeconomic situation. In response, it progressively reduced total government spending, from 30 percent of gross domestic product in 1984 to 23 percent in 1989—a very significant reduction by international standards.

At the same time, the government targeted its spending much more tightly toward the most disadvantaged in society. It was keen to obtain greater value for money from government spending, and with this goal in mind it introduced a series of innovative public sector reforms, particularly in the areas of financial management and budgetary reform. Collectively, these reforms placed Australia at the forefront of OECD countries in terms of public sector management.¹

The reforms provided much greater autonomy to the heads of line departments to manage their budget appropriations—under a philosophy of letting the managers manage. Although it had initially been hoped that the reforms would encourage departments to closely manage and measure their performance, this expectation was not met. With the support of other central departments, the powerful Department of Finance (DoF) therefore developed a whole-of-government evaluation strategy. This strategy received the cabinet's strong endorsement through a formal cabinet decision. The strategy followed the philosophy of making the managers manage.

Main Components of the M&E System²

The Australian M&E system largely comprised a formal strategy for evaluations. It was based on a 1988 diagnostic review of evaluation practices in departments and of the overall level of evaluation activity in government. The strategy itself was progressively developed over several years (1987–1991). The strategy had three principal objectives: to encourage program managers to use evaluation to improve their programs' performance; to aid the cabinet's decision making and prioritization, particularly in the annual budget process, when a large number of competing proposals are advocated by individual ministers; and to strengthen accountability in a devolved environment by providing formal evidence of program managers' oversight and management of program resources.³

A centerpiece of the strategy was evaluation planning. This was done through formal portfolio evaluation plans (PEPs), which had to be submitted annually to the minister for finance (see box 8.1). These PEPs had a rolling, three-year coverage and indicated which programs or subprograms would

Box 8.1: Australia's Whole-of-Government M&E System

The main components of the M&E system were as follows:

Formal evaluation planning through PEPs, which listed the government programs the ministry intended to evaluate and the issues to be addressed in each evaluation. These PEPs were prepared annually, on a rolling three-year basis. They included major evaluations only, that is, evaluations of programs considered strategically important to the government: programs with large budgets; those of particular policy importance; problem programs; and pilot programs. The evaluations were conducted by the line ministry itself, usually with some sort of involvement by the finance ministry.

Requirement for every program to be evaluated at least once every three to five years. In practice, this meant that some aspects (such as particular subprograms) of each program were evaluated; most evaluations did not attempt to comprehensively address all aspects of a program's performance. In addition, there was also a requirement that all completed evaluations be published, with the exception of those with national security or industrial relations sensitivity.

Reviews of each ministry's program objectives and performance reporting. These reviews were conducted jointly by each ministry and the finance department, on a rolling basis over a three-year period.

be evaluated, which aspects evaluated, and when. There was a formal requirement that every program be evaluated every 3–5 years. These PEP evaluations were classified as major evaluations. Departments were also expected to initiate other, smaller evaluations, purely for internal management purposes. By the mid-1990s, about 160 PEP evaluations were under way at any one time.

The key issues for the PEPs were the choice of programs to be evaluated and the specific questions each evaluation would address; thus, the ToRs for each evaluation were crucial. These issues were decided through negotiations between the line departments and the DoF's budget sections. For the weaker departments, finance's priorities would largely prevail. For more powerful line departments, the balance of power was more even. Unresolved disputes concerning evaluation priorities would be escalated to the level of ministers or even to the cabinet if agreement could not be reached.

The line departments were responsible not only for evaluation planning, but also for the conduct of evaluations of their programs. Some of the larger departments—such as the employment department—had an evaluation branch with some 20–25 staff responsible for planning evaluations,

providing advice on evaluation methodology, participating in evaluation steering committees, and conducting major evaluations.

Other departments had only small evaluation units, as part of a planning/coordination branch, and devolved the evaluation function to line program areas; they in turn would be responsible to the top management of the department for the quality and rigor of the evaluation (especially for smaller evaluations, which were often in the nature of rapid reviews). A number of evaluations were contracted out to individual consultants or consulting firms.

The DoF's budget sections also involved themselves in the conduct of individual evaluations wherever possible. This would often involve membership of interdepartmental steering committees for the major evaluations; these committees would usually include other relevant sector ministries and other powerful central departments such as the treasury and the department of the prime minister and the cabinet. This broad membership would allow the budget officials (1) to seek to influence the conduct of evaluations, to ensure that problems were fully investigated in an impartial manner, and (2) to comment on draft evaluation reports.

One feature of Australia's M&E system is that a broad definition of evaluation was used. Evaluation was defined as a form of disciplined inquiry: it included rapid evaluations, formal policy reviews, rigorous impact evaluations, and performance audits conducted by the national audit office. The cost of these evaluations varied widely: a sample of evaluations analyzed by the finance department ranged in cost (in 1993 prices) from about \$43,000 to \$430,000.⁴

The evaluation strategy was strengthened in 1991 in response to a performance audit that found that departments varied in their level of commitment to evaluation. The audit office report criticized some departments for poor choice of programs evaluated and for an insufficient focus on the effectiveness of government programs. In response, the DoF created a specialist evaluation unit responsible for providing advice, support, training, and encouragement to other departments, as well as to the budget areas of the DoF itself. This unit also monitored departments' evaluation planning and the number of evaluations conducted; the head of the finance department used this information to informally pressure line departments to improve their evaluation activities.

Australia's M&E system essentially stressed *evaluation*, which was viewed as providing the necessary in-depth, reliable information on the efficiency and effectiveness of government programs. Performance information was also understood to be important, but it was viewed as an issue for line departments to manage. By the mid-1990s, however, the finance department was concerned about the quality of this information and commissioned reviews of departments' annual reports and their budget documentation (which is tabled in the Parliament).

The deficiencies in these reports led the finance department in 1995 to mandate a rolling series of detailed reviews of each department's program objectives and performance information. These reviews were conducted jointly by the finance department and the line department, and recommendations for improvement were required to be implemented.

Performance Budgeting

While Australia's M&E system had three stated objectives, from the perspective of the finance department—which was the primary architect and the overseer of the M&E system—the objective to which it devoted most attention was to support the cabinet's decision making during the budget process. The DoF played a highly influential role in the budget process in Australia. It prepared policy analyses of all new spending proposals prepared by sector ministries, and these analyses accompanied the spending proposals sent to cabinet ministers for their collective consideration when making budget decisions. The DoF thus provided an independent policy analysis that typically constituted a view counter to that of the spending ministry. The work of DoF's budget sections also included the preparation of “savings options”: policy proposals to cut or abolish existing government programs.⁵

The budget process entailed a “marketplace of ideas.” In this inherently adversarial situation, having evaluation findings available about the performance of programs was an important means of ensuring a reliable evidentiary basis for budget decisions. Evaluations had the potential to provide a competitive advantage to those who relied on them. Thus it was important that the DoF's budget sections were fully involved in the evaluation planning of ministries and in the conduct of major evaluations. This ensured that the DoF budget officials were familiar with the quality and any limitations of the evaluations, were fully aware of the evaluation findings and recommendations, and were thus able to use them in their policy analysis work. Involvement of these officials in the evaluations would also increase their knowledge of the program's objectives and the realities of its operating environment—this understanding is important for their work.

It is worth noting here that finance ministries may not always be supporters of reforms designed to strengthen the amount of information available on government performance. Before the reforms, the Australian DoF was heavily involved in the detailed scrutiny of departments' spending activities. The danger is that this traditional focus on

spending can mean that relatively little attention is paid to the results of that spending. And powerful finance ministries can even act as roadblocks to reform. Having the DoF responsible for evaluation oversight ensured that there was a direct influence on the divisional units within DoF that had oversight for line departments.

However, achieving the needed cultural change within the DoF was a slow process over a number of years and involved a substantial staff turnover. The DoF's greater focus on issues of value for money (rather than on spending issues) flowed through to the nature and quality of policy advice the DoF provided to the cabinet; that advice increasingly drew on available evaluation findings.

The most important feature of Australia's M&E system is the significant use made of evaluation findings to support the cabinet's budget decision making. The DoF conducted surveys of its budget staff each year to ascertain the influence of eval-

uations on each new policy proposal of line ministers and the savings options proposed by DoF (or by line ministers if they wished to fund new policies).⁶ By 1994–95, about \$1.75 billion (or 77 percent) of new policy proposals were judged to have been influenced by the findings of an evaluation, and in most cases the influence was judged to be direct. The corresponding figures for savings options were \$380 million (65 percent of the total).

Evaluation findings influenced not only the policy options put forward for the cabinet's consideration, but also the cabinet's decisions. DoF budget officials were also surveyed regarding the extent to which evaluation had influenced the cabinet's decisions in the 1993–94 and 1994–95 budgets. The evidence is mixed, but it indicates that evaluation played a substantive role. In 1994–95, evaluation was judged to have influenced the cabinet's decision in 68 percent of the \$2.846 billion of proposals considered (new policy proposals plus savings options).

Box 8.2: Influential Evaluations in Australia

In the 1996–97 budget the new government was determined to both reduce and reprioritize government spending. Particular focus was given to labor market and related programs, which accounted for \$2.90 billion in spending annually. 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, such as achieving a better match of labor supply and demand. At the same time, the minister 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 some 6 months after he or she had participated in a labor market program. (The base-

line was a matched comparison group of individuals who did not participate in a program.)

Evaluation findings showed that the JobStart program, which provided wage subsidies, had a net cost of \$3,700 per additional job placement, whereas the JobSkills program, which was a direct job-creation program, had a corresponding net cost of \$57,800. The minister noted, "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 and the JobSkills program was substantially scaled back and more tightly targeted to job seekers who were particularly disadvantaged.

Total savings to the government from its reduction and reprioritization of labor market programs were about \$1.14 billion over two years. The 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: Mackay 1998a.

Table 8.1: Strengths and Weaknesses of Australia's M&E System

Strengths	Weaknesses
<ul style="list-style-type: none"> • Evaluation findings heavily used for budget analysis, policy advice, and by the cabinet in its budget decision making • High utilization of evaluation findings by sector departments and agencies • Evaluation conducted as a collaborative endeavor between finance department, other central departments, and sector departments 	<ul style="list-style-type: none"> • Uneven quality of evaluations • Insufficient availability of advanced evaluation training • Insufficient attention to regular performance information • A claimed administrative burden on departments

Source: Mackay 1998a, 2004.

The corresponding proportion for the 1993–94 budget, however, was only 19 percent of proposals. One important reason for this difference was the substantial revision of labor market, industry, regional, and aboriginal policies in the 1994–95 budget—the major policy review on which these decisions were based had been heavily influenced by a number of evaluations commissioned specifically to help guide the policy review.

Only a few programs were terminated as a result of an evaluation; given the emphasis in the budget process on portfolio (defined as a line department plus outrider agencies) budget envelopes—in effect, portfolio budgeting—any program termination would often result in spending reallocation within that portfolio. There were major instances where programs were significantly cut as part of a major reprioritization, for example, in the labor market and social security areas; these cuts reflected the desire to maximize value for money given the government's policy priorities (box 8.2).

Strengths and Weaknesses of Australia's M&E System

Strengths

Some commentators have observed that “program evaluation (in Australia) has been applied more extensively and systematically than in any other country” (Kim and others 2006). The outstanding feature of Australia's M&E system was the

high utilization of evaluation findings in the budget process, as a key input for both high-quality policy advice (Uhr and Mackay 1996) and for the cabinet's budget decision making (table 8.1). As noted in the chapter 5 discussion on performance budgeting, the influence of evaluation findings on the government's decisions was typically indirect—ranging from a major influence in a number of cases to little or none in others. This is essentially the most that can be expected from a whole-of-government M&E system, in terms of potential utilization of the M&E information that the system produces.

The Australian auditor-general has commented, “In my view, the success of evaluation at the federal level of government . . . was largely due to its full integration into the budget processes. Where there was a resource commitment, some form of evaluation was necessary to provide justification for virtually all budget bids” (Barrett 2001).

Another strength of the M&E system was the high level of utilization of the information by line departments and agencies. A performance audit conducted by the Australian National Audit Office (ANAO) in 1997 concluded that line departments were making considerable use of their evaluation findings to help improve operational efficiency. To a lesser extent, they were also using these findings to help guide their own resource allocation decisions and in the design of service

quality improvements (Australian National Audit Office 1997).

This high level of utilization by line departments reflected another strength of the Australian M&E system: evaluation was essentially a collaborative effort involving the DoF, other central departments, and the line departments. Although responsibility for evaluation was largely devolved to line departments, the involvement of the central departments in the planning and oversight of the major PEP evaluations helped achieve broad ownership of the evaluations themselves and of their findings.

This approach stands in contrast to that of Chile, for example, where the line ministries generally have little or no ownership of the evaluations the finance ministry commissions. The Australian approach also ensured that line departments' deep knowledge and understanding of their own programs were used intensively in evaluations; a danger with an evaluation system that relies on externally conducted, independent evaluations is that it can fail to draw on this program expertise.

Weaknesses

The downside of Australia's more collaborative approach, however, was an uneven quality of evaluations conducted by line departments. The ANAO performance audit analyzed a sample of evaluation reports and concluded that more than one-third suffered from methodological weaknesses of one kind or another.

One reason for this was that many of the program areas of line departments that had responsibility for conducting or commissioning evaluations lacked sufficient skills to ensure that high-quality evaluations were conducted. The DoF provided basic training in evaluation methods and issued handbooks on program evaluation and cost-benefit analysis. But the ANAO audit reported that 20 percent of line departments were concerned about the lack of available training in advanced evaluation techniques.

In retrospect, one option to address this issue of evaluation quality would have been for the DoF

to mandate creating sizeable central evaluation units in each department. Another option would have been to centralize the entire evaluation function in the DoF—this would have required the creation of a very large evaluation office and would have been contrary to the devolutionary nature of most of the public sector reforms.

Two other weaknesses of the Australian M&E system are worth noting. First is the relatively weak emphasis given to the regular collection and use of performance information. The DoF explicitly advised departments of the importance of developing sound program objectives and having sound performance information—not least to facilitate the conduct of evaluations. It was not until 1993, some six years after the evaluation strategy was initiated, that the DoF commissioned broad reviews of the quality of departments' annual reports and budget documentation.

Following these reviews, which were critical of the quality of program objectives and performance information, the DoF mandated a series of in-depth, rolling reviews of all departments' and agencies' program objectives and performance information. It required that action plans be implemented to address any problems identified. One lesson from this experience is that all M&E systems require ongoing review and adjustment; it is necessary to monitor and evaluate an M&E system, just as it is necessary to monitor and evaluate any other type of public sector reform.

A final weakness of the M&E system claimed by some departmental secretaries was that the formal requirements for evaluation planning and reporting were too burdensome. Some departments did create complex internal processes for evaluation planning and ended up preparing 120-page PEPs; yet the DoF guidelines did not mandate such complexity, and the guidelines recommended much smaller PEPs as good practice.

The Evaluation System—A Postscript⁷

The election of a conservative government in 1996 led to the considerable downsizing of the civil service; a substantial weakening of the policy advising process and the role of the DoF, especially

in the budget cycle; and the dismantling of many of the remaining central controls and requirements. At the same time, a much higher level of autonomy was given to the heads of line departments—a return to the philosophy of letting managers manage.

Although line departments are still required to report their performance publicly to the Parliament, the DoF now applies no central controls or quality standards for M&E. As part of these changes, the decade-long formal evaluation strategy—and thus the government’s M&E system—was dismantled. The Australian auditor-general has characterized these reforms as a deregulation of evaluation. While there remain some line departments that can be considered to be good practice “islands” of M&E,⁸ in terms of their conduct, quality, and use of M&E, these appear to

be the exception rather than the rule.⁹ ANAO performance audits have highlighted the poor quality of performance information that most departments now provide to the Parliament.

Although the DoF still provides advice on departments’ budgets, it lacks systematic, reliable monitoring information and evaluation findings on which to base this advice. As the OECD has concluded—

In Australia, the deregulation of the public service and the adoption of an arm’s-length posture by the central agencies allowed management freedom but is currently considered to have deprived the Finance Ministry of the information necessary for it to adequately advise the Minister (OECD 2002).