

‘Death by evaluation’? Reflections on monitoring and evaluation in Australia and New Zealand¹

mandatory evaluation

This paper reviews the state of play in outcome-oriented monitoring and evaluation in New Zealand in the light of recent Australian experience. Given issues arising from a ‘big bang’ approach to evaluation adopted with the Commonwealth of Australia’s 1988 Evaluation Strategy, and the lead times involved in building capacity, it argues that New Zealand should focus in the short-term on creating ‘evaluative management cultures’ inside organisations rather than opting exclusively for building formal and technical evaluation capacity and expectations – although that too should be developed, even though it will take some years to mature. Actually, a practice rather than institutional approach to capacity building that tries to ‘build from within’ would also be consistent with current thinking in management and organisational development. More may be gained by developing a learning approach to public management, strategic emphasis on internal evaluation, and focusing on immediate and intermediate outcomes especially in relation to monitoring.

Some in Aotearoa/New Zealand believe that the Commonwealth suffered ‘death by evaluation’ during the 1990s. Considering the truth or otherwise of this assertion has led me to certain proposals about where this country should head in coming years regarding performance monitoring and policy and program evaluation. This paper works through the arguments.

Some evaluators may find the proposals controversial. Public managers, however, are less likely to do so, if only because they bridge the gap between ‘evaluation’ and ‘management’ and are driven by pragmatic and contextual concerns. In overview, I argue that:

- As is now widely accepted, monitoring and evaluation in New Zealand agencies need to be significantly upgraded, partly to improve the quality of management practices, partly to refocus public management on the outcomes being achieved, and partly to improve the

Bill Ryan



*Bill Ryan is Associate
Professor and Director
of Programs in the
School of Government,
Victoria University of
Wellington, Wellington.
He can be contacted at
<Bill.Ryan@vuw.ac.nz>.²*

flow of performance information and accountability between the executive and the parliament.

- However, the preconditions do not exist in New Zealand for prescribing a rapid and intensive introduction of performance monitoring and program evaluation as attempted by the Commonwealth with its 1988 Evaluation Strategy.
- Instead, New Zealand should adopt a two-pronged strategy designed to deepen practice over several years. A start should be made on building formal requirements and expert capacity for evaluation – recognising, however, that this will take time to produce results. In the short term, consistent with current thinking in management and organisational theory, more emphasis should be placed on building evaluative cultures and practices from the ground up amongst program staff, managers, agencies and providers, developing internal, learning-oriented evaluation, and refocusing monitoring activities on immediate and intermediate program and policy outcomes.

The Commonwealth: ‘death by evaluation’?

As noted above, the view from this side of the Tasman seems to be that (a) Australia now has a thriving evaluation industry but (b) the Commonwealth Government went overboard with its 1988 Evaluation Strategy. It is asserted that this strategy demanded too much evaluation too quickly, that evaluation requirements came to dominate everything else, and that a compliance culture in evaluation emerged as a result. The lesson drawn in New Zealand, where the lack of evaluation is now widely acknowledged, is that the ‘big bang’ approach should be avoided. There is probably some validity in these views.

What was the 1988 Evaluation Strategy? How did it come about? What happened as a consequence?

Barrett (1992) describes how, after three years of the Financial Management Improvement Program, the (then) Commonwealth Department of Finance (DoF) concluded that the level of evaluation being conducted within portfolios was too low and the quality too uneven. In 1988, the Commonwealth Government announced the Evaluation Strategy for immediate implementation. This entailed principally:

- the development of five-year Portfolio Evaluation Plans (including budget allocations)
- all major programs subjected to five-year rolling evaluations
- a bias in favour of publication of the results.
- the extensive provision of outcome-oriented monitoring and evaluation information in the agency’s Program Performance Report. This was

a reporting document, designed to sit alongside the Annual Report, which was supposed to provide an extensive array of outcome-oriented performance information to parliament as part of the Estimates process.

It is also worth noting that the approach to evaluation promoted by the DoF was decidedly positivist and technocratic, and hence evaluation methodologies and methods were deemed to be relatively straightforward and needed only to be adopted to produce the desired results. These proposals, of course, were problematic (see various papers in O’Faircheallaigh and Ryan 1992).

Some State governments also increased their commitment to evaluation around the same period, and by the early 1990s, evaluation activity throughout Australia seemed to be on the increase.³ Issues, however, were becoming apparent, including:

- the supply of sufficient skilled evaluators
- the long lead time required to build evaluative capacity across all agencies
- difficulties in obtaining and ring-fencing evaluation budgets and preserving the independence of internal evaluation units
- understanding the complementarities of quantitative and qualitative methodologies and having the legitimacy of the latter accepted
- the availability and integrity of long-run qualitative and quantitative data sets
- the need for a mature and outcome-oriented approach to performance by senior management, ministers and the parliamentarians, especially in the estimates process.

In addition, the intellectual difficulty of, and the amount of work involved in undertaking, high-quality performance monitoring and program evaluation caught newcomers by surprise. It was not long before the Commonwealth portfolios were looking to reduce the compliance pressure through a more selective and strategic approach to program evaluation and a reduction in their reporting load. By 1993–94 some of the formal requirements were de facto being wound back to what were probably more realistic levels. The principles of the Evaluation Strategy were maintained, but stretched out on a longer implementation timeline, and at a lower level of prescription.

Unfortunately, the new Coalition government came to power in 1996 with a different public management agenda based on economic rather than managerial reform, similar to New Zealand’s widely applauded model of the late 1980s. This was introduced with missionary zeal – as it had been in

As noted above, the view from this side of the Tasman seems to be that (a) Australia now has a thriving evaluation industry but (b) the Commonwealth Government went overboard with its 1988 Evaluation Strategy.

this country – and, convinced of its correctness, Canberra seemed to lose interest in evaluating the actual outcomes of these changes or the effects of its policies. The rise and rise of evaluation faltered, kept alive only in policy arenas where evaluation was regarded as good professional and management practice, including education, social policy, community services, criminology, family studies, regional development and labour market programs.

So what was/is the situation in the Commonwealth at the beginning of the 21st century? From this side of the Tasman it appears that a fair amount of monitoring and evaluation is being carried out, some good, some poor, some strategic and some ad hoc, but that overall, the quantity and quality are not as good as they might have been. Capacity seems to again be improving, particularly amongst external evaluators in the private, community and public (tertiary) sectors, but the internal capacity of agencies to understand, design and manage monitoring and evaluation has not improved at the same rate. On the other hand, public management in the Commonwealth is generally outcome-oriented as well as output-oriented and, as indicated by the types and content of plans and reports on many Commonwealth government websites, there are signs of an ‘evaluative management culture’ – an evaluative approach embedded in agencies motivated more by the ongoing search for substantive policy improvement than by the need or desire to comply with formal accountability prescriptions. In this respect the contrast with New Zealand, where few such developments have taken place, is marked.

Overall, therefore, the following conclusions can be drawn:

- The 1988 Commonwealth Evaluation Strategy *did* expect too much too soon. Moreover, evaluation was regarded as a straightforward management tool that only had to be used to start producing results. The immediate effect was evaluation overload, patchy quality, unselective evaluation, and a certain waste of money. ‘Death by evaluation’ is an applicable – if tart – judgement.
- But the positives achieved in the Commonwealth should not be ignored. An evaluative management culture now permeates some agencies and there seems to be a level of experienced evaluation capacity. This positions the Commonwealth well in relation to current trends in Australasian-Anglo-American public management towards ‘managing for outcomes’.

Monitoring and evaluation in the New Zealand model of public management

Criticisms have been made for several years of the low level and poor quality of performance monitoring and program evaluation in New Zealand (e.g. Boston et al. 1996; Schick 1996); this country seems to be alone amongst the liberal

democracies in not progressing public management in this respect. Actually, no-one should have been surprised, since this outcome was predictable (Ryan 1993). The New Zealand model of public management flowed out of an economic reform agenda, rather than a managerial one, that adopted outputs as the basic unit of public management, not only for budgetary and financial management but also for program and policy management (e.g. NZ Treasury 1987). The logic of public management cast in this mould prioritised economy and efficiency, with accountability focused on the control of agents (public managers, providers) by principals (ministers, funders) through the Output Purchase Agreement and the Chief Executive’s Performance Agreement.⁴ Performance in producing outputs, goods and services – the things agencies said they would produce and the activities they said they would carry out – was made the basis of accountability; an approach that militated against the eventual development of outcome-oriented performance monitoring and program evaluation.

The core features of the so-called ‘New Zealand model of public management’, including the roles of the Treasury and the State Services Commission (SSC) in managing and developing the new system⁵, were embedded in the *Public Finance Act 1989* and the *State Sector Act 1988*. Notwithstanding:

- Some agencies had a historic interest in outcomes and evaluation. Examples include the Department of Labour (DoL), the Ministry of Social Development (MSD, previously the Ministry of Social Policy), and the Department of Child, Youth and Families (DCYF), all of which undertake systematic, albeit limited, programs of evaluation and publish the results.
- Following the Schick review (1996) and the recent Ministerial Advisory Group Review of the Centre (Ministerial Advisory Group 2001), shortcomings in relation to monitoring and evaluation were increasingly acknowledged, driven partly by the current Labour government asking unambiguous questions about the results being achieved by past and present policy and not being too satisfied with the type of answers provided. Current projects taking these developments forward include *Pathfinder* and the *Statements of Intent Roll-out*.⁶
- Treasury now has a limited requirement for evaluation in relation to new budget bids. Ministers and agencies must now indicate how and when interventions proposed will be evaluated.
- As demand slowly increases, supply is also expanding with a small number of evaluators and evaluation centres and companies starting to emerge and consolidate – although the lack as yet of a national training program in New Zealand is barrier to building capacity.

In this context, it might be tempting to argue that New Zealand should adopt a ‘big bang’ strategy, as the Commonwealth did in 1988, to force

the growth of evaluation capacity and practice. But that would probably not be wise. By the mid- to late 1980s, public management in Australia had developed some of the supply and demand factors that allowed some growth in evaluation; even so, the 1988 Evaluation Strategy tried to do too much, too quickly. New Zealand does not yet have even the preconditions. In addition, it is now confronting the same challenges as other jurisdictions in relation to the cost and labour intensity of evaluation, the availability of skilled and experienced evaluators, and so on. Complicating the picture further, there are paradigmatic movements occurring in evaluation internationally that make conventional solutions problematic. A more thoughtful and contextually relevant response is required.

The rest of this paper deals with some of these issues. It concludes that a dual strategy is required in New Zealand to expand evaluative capacity and practice. One half of that strategy is conventional and, for the purposes of this paper, is assumed and little discussed – reappearing in the final section of this paper. In brief: since evaluation and monitoring are inherently technical activities, there is an obvious case for increasing the provision of specialist education and training for evaluators, for recruiting applicants into those programs and streaming them into the areas of greatest market demand inside and around the public sector. Having said that, given international developments in evaluation, this training must include positivist, non-positivist and post-positivist approaches to evaluation (Ryan 2003) and have a significant practical component so that participants are job-ready when they complete their training.

The following sections of this paper deal with the arguments underpinning the other half of the proposed strategy – an approach to capacity- and practice-building in monitoring and evaluation that is not quite so conventional but is more likely to

generate useful developments in the short- and medium-term.

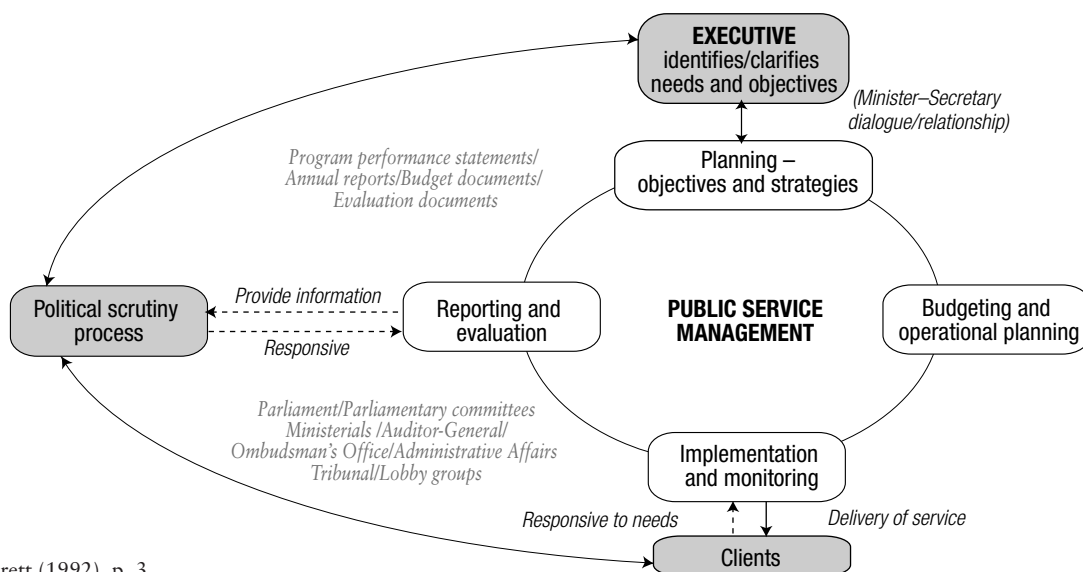
Creating an evaluative management culture

By ‘evaluative management culture’ I mean an organisational culture wherein policy and program managers in the development and implementation stages constantly check whether their efforts are having a positive (or any) effect for clients and society (presumably, although not necessarily, as specified in the program goals and objectives) and, if not, what needs to be done instead. The term refers to the beliefs and values, norms and routines embedded in an organisation that underpin the practices of performance monitoring and policy and program evaluation. It also registers that monitoring and evaluation occur in many different ways, formal and informal, as part of tacit thought and action and accumulated experience inside the organisation, whether or not the processes and procedures have been codified and elaborated, and whether or not the organisation has the financial and human capacity to undertake extensive, formal evaluation.⁷

Those organisations most likely to develop such a culture are probably those that are also committed to a client focus, strategic public management and the achievement of policy outcomes. In that sense, ongoing monitoring and evaluation – formal or informal – are inseparable from the management cycle – a representation of the management process that is familiar to Australian public managers. Figure 1 is typical (from a 1992 DoF perspective) of how the strategy/evaluation relationship is expressed in a wide range of strategic management literature predicated on any kind of ‘management by objectives’ approach.

Public management in Australia, and especially strategy and evaluation, is founded on this model more than in New Zealand. In this country, public

FIGURE 1: THE PROGRAM MANAGEMENT CYCLE



Source: Barrett (1992), p. 3

management has been focused primarily on the internal management of organisations rather than management of outcomes, and on the production of the outputs specified in the agency Purchase Agreement. Meeting the terms of this agreement has defined the meaning of 'performance'. Accordingly, monitoring has been largely process-oriented, likewise evaluation, whereas strategic and evaluative public management should be more concerned with whether the organisation is achieving those policy purposes for which it was set up. This general observation is valid even though isolated pockets of practice in New Zealand indicate that, in places, strategy and evaluation have developed along more promising lines.

In fact, the state of play in New Zealand regarding monitoring and evaluation might have become better than it is. The Strategic Result Area/Key Result Area system⁸ instigated by the National government of the 1990s (Boston & Pallot 1997; see also SSC 1998) held considerable promise in this respect, but was discarded by the present Labour-led coalition on the grounds that it was 'policy of the previous government'. Further, despite its recommendation by the Review of the Centre, the

current government is reluctant to engage in the 'strategic conversation' required between ministers and senior managers in order to disassemble abstract policy aims and priorities to identify cross-agency, departmental and program goals and objectives so that

... the current government is reluctant to engage in the 'strategic conversation' required between ministers and senior managers in order to disassemble abstract policy aims and priorities to identify cross-agency, departmental and program goals and objectives so that 'managing for outcomes' can develop and mature.

'managing for outcomes' can develop and mature. Neither strategy nor evaluation is helped in the process.

What other elements might comprise an 'evaluative management culture'? Four spring to mind. They are:

- an outcomes/strategy/effectiveness orientation to public management
- a learning approach to management where monitoring and evaluation provide ongoing information
- an emphasis on internal monitoring and evaluation
- development of (immediate and intermediate) outcomes monitoring.

An outcomes/strategy/effectiveness orientation

It is difficult to imagine how an evaluative management culture could exist outside an 'outcomes' orientation. In New Zealand, the recommendations of the Schick report (1996) and now the *Review of the Centre* (Ministerial Advisory

Group 2001) have put outcomes at the forefront of the public management agenda. The hard separation between ministers and public servants of task and accountability prescribed by the *Public Finance Act 1989* – the former responsible for outcomes and the latter for outputs – was a practical fiction aimed at simplifying accountability arrangements that, at best, made heroic assumptions about (a) the purchasing capabilities of ministers, and (b) the capability of public servants to elaborate the hypothetical links between outputs and outcomes; that is, ex-ante specification of what in New Zealand is called the 'intervention logic' (the Australian equivalent is 'program logic'). The practical effect of the New Zealand preoccupation was that agencies had no motivation for focusing on policy and program outcomes – which, it can be argued, denies the very purpose for which public services emerged in liberal democracies in the first place. Now, pressing questions are being asked by policy communities and government about the policy outcomes being achieved by government agencies, and several projects are now under way designed to move the public sector progressively in that direction. Importantly, in December 2001, Cabinet signed off on the introduction of 'Managing for Outcomes' (MFO). MFO is an outcome-based approach to departments' planning, management and reporting, the aim of which is to improve the performance of the public service, requiring departments to adopt a strategic and outcome-focused approach to planning, management and reporting while focusing on delivering outputs.

Because the Ministerial Advisory Group (2001) argued that the New Zealand model of public management paid insufficient attention to outcomes or the connections between particular outputs and outcomes, the Implementation Group, working out of the SSC but including the Treasury, Department of Prime Minister and Cabinet and Te Puni Kokiri (Department of Māori Development), are taking the recommendations of the Review forward. These include the introduction of Statements of Intent to be produced by agencies as part of the budget process (Robinson 2002; also described briefly in note 6). The SSC itself is attempting to model the desired new ways of acting, partly by restructuring itself around its new Deputy Commissioner Teams whose task is to help agencies improve the strategic qualities of their internal and external management.⁹ In fact, by the turn into the new millennium, driven partly by their own frustrations with the output-based approach to public management, some agencies had already started moving towards MFO without waiting for official modifications to the model or the relevant legislation. Examples are the Ministry of Health (e.g. primary health care strategy, disability strategy), the Ministry of Education (e.g. adult literacy), the Ministry of Economic Development (e.g. regional development), the Ministry of Social Development (e.g. Strengthening Families) and the Ministry for Conservation (Environment 2010: Biodiversity Strategy)(see also Ussher & Kibblewhite 2001).

The important thing here is that there is a logic underpinning MFO and its concern with effectiveness that leads more or less inevitably towards an evaluative management culture. If managers and staff are focused primarily on achieving outcomes, and are receiving signals from ministers and parliament that their effectiveness, responsibility and accountability will be defined in those terms, then a reflexive evaluative orientation will feed back into the practical routines that constitute program and policy management.¹⁰ This has already occurred to a considerable extent in Australian public services and in some pockets within some New Zealand agencies. The implementation trick will be to build on what they have already achieved and to generalise it throughout all agencies.

A learning approach to management and the role of monitoring and evaluation

The emergence of a 'learning' approach to management with its recognition of 'emergent' strategies and outcomes is one of the major shifts in public management of the present era. It signals important changes in the approaches, methods and tools of management and acceptance of new organisational realities. It also seems likely to become part of the public management orthodoxy of the early 21st century (e.g. OECD 2000, 2001; see also Giddens 1998).

Despite their impact in comparable jurisdictions (Britain's Blair government has explicitly adopted them, e.g. UK Government 1999; as has the Canadian government, e.g. CCMD 2001), they have barely penetrated into New Zealand public management. In fact, the culture inside New Zealand agencies remains remarkably hierarchical and risk-averse. Relationships between the present cabinet and agencies, between central and line agencies, and within the agencies themselves, are centrist and biased towards command and control. In this context, performance monitoring and program evaluation have been applied as instruments of control; practices to be feared by line agencies for the sanctions they may bring. Given the role of public choice and principal/agent theory in shaping the New Zealand model of public management, this is probably not surprising.

But the 21st century context of increasing complexity, uncertainty and discontinuity demands new forms of governance and management (e.g. Giddens 1990; OECD 2000; 2001). Control from the centre, technocratic solutions to policy problems, the desiderata of rationality, tractability and linearity in policy design, are now regarded as inapt. Instead, subsidiarity (more commonly described in this part of the world as 'devolution and localism') is regarded as essential, as are flexibility and responsiveness in implementation and delivery. So are new forms of engagement between citizens and state. Policy and strategy are now often regarded as hypotheses or visions based on

uncertain and incomplete knowledge, undertaken in difficult and puzzling circumstances that often give rise to inexplicable effects, against which actual outcome evidence must be constantly collected and analysed (e.g. Ryan 2002). Reconstituted by these trends, public management in the liberal democracies in the 21st century is already becoming highly contingent in character, constantly conditional, wherein the ongoing need for ongoing organisational, advocacy, policy and management learning is crucial.

Such circumstances make monitoring and evaluation, the constant review of actual outcomes and their comparison with the original intentions, even more vital. Further, the recursive nature of action and understanding often leads to redefinition of the original goals and objectives in the light of new developments; the 'emergent' character of progress may be the most radical aspect of the learning paradigm and provides the greatest challenge to conventional political and policy thinking in Westminster-derived polities. It undermines formalised notions of 'performance', where ex-ante specification of goals and objectives is traditionally taken as a political promise against which the executive must perform or be held to account by the parliamentary opposition.

No 'official' model of public management in Australia, New Zealand, the US, the UK or Canada as yet fully reflects these developments. But interesting – and often unsung – experiments are taking place in various sites in Australia and New Zealand (e.g. domestic violence reduction programs in both countries; similarly for many community and regional development and coordinated service delivery initiatives) and, as such, they show us something of the future. Monitoring and evaluation as the foundations of management and policy learning are inherent to the development and implementation of these initiatives, as they must be for any approach to public management that takes seriously collective policy and organisational learning.

A strategic emphasis on internal evaluation

The classic argument is that evaluation should be carried out by independent, usually external professionals exercising the highest levels of methodological rigour and producing data and generalisations that, as far as possible, make a justifiable truth claim. From proximately 'correct' knowledge, so the argument goes, effective policy (re)development can proceed.

As all evaluators are aware, the 'scientific' approach to evaluation, predicated on one or another positivist epistemology and applied in open, social and community settings, has undergone significant challenge in recent decades with a collection of non-positivist approaches (often wrongly called 'qualitative'). This includes, notably, the so-called 'Fourth Generation' approaches now well established as part of the evaluation stock-in-trade (for an overview see Ryan 1996). As

positivism was undermined and a new political, value and methodological relativism overtook social research throughout the 20th century, managerial criticisms in very recent decades relating to the cost, relevance and timeliness of evaluation also gathered strength – to say nothing of the alarm expressed during the 1970s as evaluators in the US and elsewhere tried to figure out whether they did anything useful. Carol Weiss (1977) consoled with her ‘enlightenment’ thesis, Michael Patton promoted ‘utilisation-focused evaluation’ (for an overview see Shadish, Cook & Leviton 1991). Others promoted client and/or stakeholder-centred evaluation and, more recently, client and/or stakeholder-based evaluation, where clients and stakeholders themselves undertake and conduct aspects of the evaluation.

In short, alternative methodologies, epistemologies and even ontologies have undercut the confidently ‘scientific’ pretensions of professional evaluation, particularly as practised in the US during the 1960s and ’70s. In contrast to its underlying behaviourism, many contemporary approaches to social research assume (in the structurationist terms used by Giddens; e.g. 1984) that the social actor (or ‘agent’) is the knowing subject of action – an expert in their setting, as revealed in their capacity to (re)construct and negotiate their social relationships and to theorise about their roles and actions (their practical and discursive consciousness). Recognition of agents’ expertise allows researchers and evaluators to treat their subjects’ knowledges and understandings as important – indeed, some of *the* most important – data available. Once these kinds of assumption are accepted, it is only a small step to regarding participants in the policy process – from ministers and policy analysts to managers and staff, clients and stakeholders in service delivery – as agents capable of and interested in the description and analysis of their engagement, and able to make important and detailed contributions to its evaluation. Unwittingly, perhaps, the so-called ‘steering group’ or ‘project group’ approach to evaluation promoted widely in Australia in the 1980s combined the benefits of independence and objectivity of external evaluators with the detailed and comprehensive knowledge and expertise of program staff, clients and stakeholders – a kind of internal/external approach to the management of evaluation (Ryan 1992). These days, organisational and policy learning make it increasingly sensible to encourage the communities of practice inside the policy community to engage in extensive critical reflection on the process and outcomes of their endeavours – which emphasises connections between notions of internal evaluation, learning and evaluative management cultures.¹¹

Writing in the early 1990s (e.g. Ryan 1992), in line with the conventional wisdom of the time wherein internal evaluation was regarded as problematic, I advocated external, expert evaluation as representing the best way forward for public organisations in Australia. Since then, practical

experience with evaluations based on the internal/external approach has led me to a different view. Some of the most searching critics I have worked with in evaluating policies and programs in Australia have been those officials involved in their development and implementation, and their contributions invariably improved the insight of the analyses and the viability of the evaluation recommendations. Other evaluators in the English-speaking world have been moving in a similar direction, towards greater enthusiasm for the possibilities inherent in internal evaluation. There are always situations where a traditional independent, external evaluation is most appropriate, but in others, an internal approach is more useful.

This approach is based on collective policy learning involving officials, providers, clients and stakeholders. The organisation and conduct of the evaluation as a whole is made the responsibility of those who are insiders to the program or policy – although technically skilled outsiders might still undertake some technical aspects of the evaluation. This kind of approach goes hand in glove with a critically reflexive form of management that trickles down across and engages all program staff and providers, focusing the attention of all on effectiveness and appropriateness. There are additional benefits to this approach: those with responsibility for implementing them own the final conclusions and recommendations, and it reduces the ‘threat’ posed by evaluation.

Heavy reliance on internal (or mixed internal/external) evaluation seems like a useful direction for New Zealand to move over the next three to four years. Entirely independent, expert evaluation can alienate the policy community from the review process. It can be expensive and time-consuming and there are few experienced evaluators on the ground to do the work. Building formal evaluation capability via the traditional education and training pathway, whether in-house or external, will take a long while. Adoption of internal evaluation within an evaluative management culture would represent more efficient use of limited resources and have significant upskilling effects. Line and central agencies can and should collaborate to generate this type of development and to do so in a justifiable way that does not fall into the twin traps of self-interest and turf protection.

Actually, a principled approach to internal evaluation potentially links to another methodological possibility – one that seems immediately attractive although fraught with possible dangers. Given the contingency of contemporary public management, managerially useful evaluation often does not require the same level of rigour or the same level of investment or the extended timeline as scientific social research – ‘managing’, after all, is often an 80/20 and even a 60/40 game. It is probable that knowledges-for-action are quite different from knowledges-as-science, hence that evaluation-for-action could afford to follow different rules. The financial and

organisational politics of evaluation impose trade-offs in cost-effectiveness, timeliness, validity and reliability, but these can be carefully managed in an internal evaluation to produce a valid, justifiable and timely evaluation, albeit one at a 60–80% level of confidence rather than 95–99%. In this context, there is a strong case for what might be called ‘rough and ready but relevant, timely and useful’ evaluation: I prefer not to use the more common phrase ‘quick and dirty’ for fairly obvious reasons – although in referring to ‘rough ‘n’ ready’ evaluation, no connotation of ‘sloppiness’ should be taken. The aim would be to produce useful knowledge, *pragmatically* and *strategically* useful knowledge, recognising also that all knowledges are relative and there are many knowledges. The possible utility of such evaluation in ongoing policy management is self-evident, even if it is difficult to identify immediately what such an approach might entail. The challenge ahead will be to find approaches and tools that enable valid, useful, timely and defensible evaluation, but which are aimed more at providing strategically useful ways forward rather than truthful accounts of what has been.

Development of (immediate and intermediate) outcome monitoring

It is widely accepted that well-designed, well-analysed, outcome-oriented monitoring information not only aids ongoing strategic policy management, but is a boon to evaluators when it comes time to undertake process and impact evaluations (e.g. Wholey & Hatry 1992). For this reason alone, New Zealand agencies would be well advised to move in this direction and away from the previous preoccupation with processes and outputs. In addition, creating an evaluative management culture from the ground up would be assisted if agencies moved into immediate and intermediate outcome monitoring. It seems to be a relatively efficient way of buttressing evaluative practices inside and between agencies, and providing significant impact for each dollar spent in doing so.

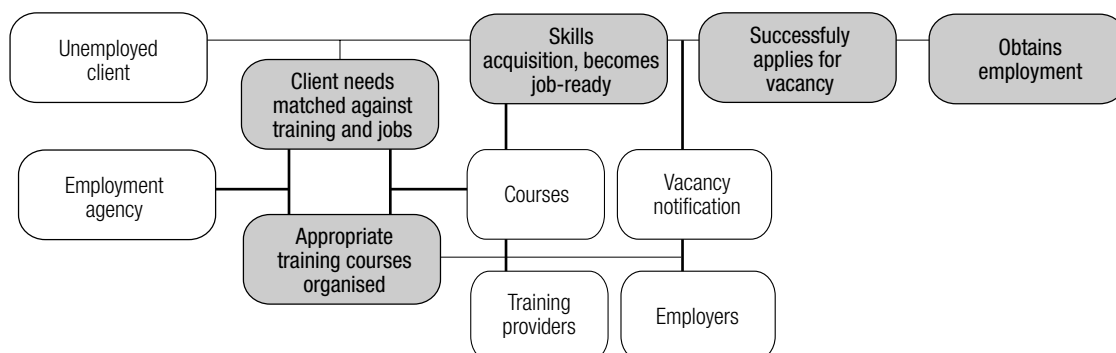
What does ‘immediate and intermediate outcome monitoring’ mean? Throughput and process data are quick and relatively easy to collect and monitor

and will always have a value for internal management purposes. However, this type of data tells program staff nothing about the impacts of implementation and service provision – the immediate outcomes of clients accessing and using agency outputs or the subsequent effects on their status at the collective or institutional levels. At the other end of the scale, societal-level monitoring of changing statuses and conditions (for a New Zealand example, see the Ministry of Social Development/Treasury indicators of social wellbeing) may say something about ultimate policy impacts, but only in unspecific, often non-attributable ways and perhaps only several years after implementation begins. Too much attention to ultimate outcomes in monitoring, therefore, is little use in improving program and policy management. To do this, staff, managers and providers need information regarding the immediate and intermediate outcomes of service delivery and implementation; the changes created in clients’ individual and collective status or their conditions of existence as a possible direct or indirect result of policy implementation and service delivery, as a result of their utilisation of the outputs produced by agencies. Outcomes that are consistent with (or acceptable in terms of) goals and objectives will tell them they are doing something right and to keep going. Those that are not will tell them to do something different.

Immediate and intermediate outcomes are the circumstances and conditions achieved with service delivery that will in some sense combine over time to generate the desired ultimate program and policy goals and objectives. The goals and objectives of implementation and service delivery can be specified ex-ante in a manner consistent with the intervention logic (whether derived conceptually from a logic-in-theory or via practical reasoning from a logic-in-use makes no theoretical difference). Indicators of these outcomes can then be developed for monitoring, and the information derived can be used to assess the ongoing effectiveness of the program and its delivery.

An illustration is presented in Figure 2: a simple intervention logic for an active labour market program, where the shaded boxes refer to outcomes.

FIGURE 2: SIMPLE INTERVENTION LOGIC FOR A CONVENTIONAL LABOUR MARKET PROGRAM



In addition to helping management develop necessary and sufficient implementation and delivery strategies, the intervention logic also provides a framework for developing sets of performance indicators (PIs). It enables the agency or provider to identify the immediate and intermediate outcomes to be achieved along the way that should be monitored as the program is implemented. Obvious examples in Figure 2 relate to stages of delivery of the program: for example the knowledge and understanding the agency and clients have of their circumstances and what needs to be done; the learning outcomes of training; the increased capability to apply for appropriate jobs, the achievement of positive employment outcomes; and so on. Combined with key (ultimate) outcome indicators, a basic set of PIs might focus on the following:

- the extent to which the agency understands the needs of its clients and is able to provide appropriate and effective levels of service, and the extent to which clients understand what they have to do to compete on the labour market and are motivated to do so
- the extent to which clients attend appropriate courses and have demonstrably expanded and increased the range and types of job-relevant skills. The attitudes, motivations and self-perceptions of clients as they progress through job-readiness training
- the extent to which clients are able to, and do, apply for a wider range of jobs and/or are unsuccessful or successful in their applications: the match of jobs available/applied for relative to the existing and new skills of the client¹²
- (assuming constant participation) increases in the proportion of people in that category who are employed and a corresponding decrease in the proportion unemployed
- (assuming constant participation) an overall rise in the employment rate and an overall reduction in the unemployment rate.

The time involved, or the intellectual difficulty in generating immediate and intermediate outcome indicators from an intervention logic, should not be underestimated, especially if this work is done collaboratively by agency and providers – even more if clients themselves are involved. But experience is already indicating the value of doing so, especially if the intervention logic is treated as a ‘heuristic’ rather than a ‘tool’ (as it should be) and this work is done (as it should be) in the policy development stage. This focuses management and staff on the desired outcomes and how evidence of actual outcomes will be collected and assessed (and what will count as evidence), and is a practical powerful driver in the creation of an evaluative management culture. It will also ensure that from the outset, the agency is collecting performance information that is useful to the provider, the funder, for reporting performance to parliament

and to the evaluators who, at a later date, will want to assess the effectiveness and appropriateness of implementation and service delivery.

A two-pronged strategy for building evaluation capacity in New Zealand

New Zealand clearly needs to significantly upgrade its approach to monitoring and evaluation as a key element of its model of public management, but should probably not do so by emulating the Commonwealth Evaluation Strategy of the late 1980s. Even assuming that a technical and expert approach to evaluation is appropriate to the conditions applying in 21st century polities, few of the requisite preconditions are met in the New Zealand context.

This analysis suggests that a two-part strategy is the most appropriate way forward for New Zealand. One would be to facilitate the creation of systematic and specialist evaluation training programs (degree and skill-based), preferably at a national level, and to encourage individuals to participate. This is an obvious strategy with which few would disagree, but will take some years to come to fruition – certainly in providing the flow of skilled labour required to match agencies’ needs.

The second strategy is to encourage forms of public management practice that build evaluative management cultures inside public sector agencies. This would probably include the following:

- A general injunction that agencies build their evaluation capacity and activities in the expectation that, within the next three to four years, a limited requirement will be legislated, particularly in relation to policy and program performance reporting to parliament.
- Establishing formal but initially limited expectations for ongoing performance monitoring and evaluation of strategically significant programs by agencies, including cross-agency collaboration in the monitoring and evaluation of ‘joined-up’ policies, and stressing the collective learning approach to underpin them (including expecting estimates and select committees to act in kind).
- Expecting cross-agency evaluations of whole-of-government policies to be undertaken as well as single programs within agencies, encouraging adoption of a range of internal strategies all focused on building an evaluative management culture within and across the public sector by making current management practices more strategic, outcome and effectiveness focused.
- The concentration of monitoring and evaluation expertise in one of the central agencies (probably the SSC¹³ or some external unit) and making it available to agencies as advisors and consultants.

This second strategy will probably bring the greatest benefits over time. Developing capacity and

undertaking evaluations will need to be funded, but the work of instigating evaluative management cultures can be commenced with committed leadership and without additional funding – since its beginnings can be developed within conventional and prescribed management practice.

But central agencies also have an important role in enabling this organisational development – the role of the SSC will be particularly important here. The new *modus operandi* being taken up by the Commission – the facilitation and brokerage of organisational and management development through the Deputy Commissioner Teams working closely with agencies over time – is, in its conception, a genuinely innovative approach to central agency work that is appropriate to the times. This approach, combined with MFO, is being implemented across the public sector over the next few years. It is to be hoped that the results will be as good as the promise. Ongoing monitoring and evaluation will enable us to know.

Notes

- 1 This is a revised version of a paper first delivered at a BIIA Conference, 'Public Sector Performance', held in Wellington, October 2001.
- 2 Bill Ryan is Associate Professor and Director of Programmes in Victoria University's School of Government in Wellington. Over the years, Bill has undertaken several evaluations for various governments and published and consulted on many aspects of public management reform and practice in Australia and New Zealand. His present research focuses on the second stage of public sector reform in New Zealand, and the development of 'managing for outcomes' on both sides of the Tasman.
- 3 The situation regarding evaluation throughout the 1990s in Australian State public services seems confused and difficult to characterise – which is why this paper focuses on the Commonwealth. A Centre for Australian Public Sector Management (CAPSM) state-of-play survey in the early 1990s (O'Faircheallaigh & Ryan 1992) suggested that States such as South Australia and New South Wales were making significant progress, but overall, development seems to have been patchy and variable. By the mid- to late 1990s, National Competition Policy and the introduction of accrual budgeting and reporting emphasised economic evaluation. Policy fields with a long history of evaluation (e.g. health, education, community services) maintained their commitment to ongoing review across all levels of government.
- 4 The Output Agreement was an annual agreement signed between the chief executive and the minister, which specified which organisational outputs the minister wanted to 'purchase' in order to achieve the outcomes sought by government. The CE's Performance Agreement was also an annual agreement based in large part on the Output Agreement plus other priorities set by the minister.
- 5 See Boston et al. (1996) for a more complete overview. It is worth emphasising that in New Zealand, no agency has taken responsibility for development of non-financial aspects of public management, as did the (then) Department of Finance and the associated Management Advisory Board/Management Improvement Advisory Committee in Australia. The State Services Commission (SSC), the equivalent to the Australian Public Service Commission, has tended to focus on the employment, management and appraisal of chief executives and some aspects of overall human resource management; the *State Sector Act 1988* devolved human resource management powers to CEs and the SSC played a minimalist role during the 1990s. Treasury, of course, played a central role in developing financial practice. The relatively underdeveloped level of non-financial management practices in New Zealand agencies now – including monitoring and evaluation – may be due in part to the low-profile approach of the SSC in this respect during the 1990s.
- 6 Pathfinder involves the SSC and Treasury working with selected agencies to help them shift their planning, delivery and reporting processes towards a conventional outcome-oriented program management framework. The Statement of Intent (SOI) is a document, produced for the annual budget process and signed off by the minister and the chief executive, that identifies policy and organisational objectives to be generated over the coming budgetary year and the strategies whereby these will be achieved – in effect, a summary of the strategic plan. All agencies were required to produce an SOI for the 2003–4 budget round.
- 7 In the same way that writers such as Mintzberg (e.g. 1994) regard 'strategic thinking' and 'strategising' as more important than 'strategic plans', so should 'evaluative management cultures' be prioritised over 'evaluations'; ultimately, the culture and approach are more significant and will contribute more to policy and program effectiveness than the production of objects or the setting up of specialist organisational units separated from policy and delivery.
- 8 In essence: government would nominate its Strategic Results Areas (SRAs) and agencies, and in dialogue with their minister(s) would identify, plan around and report on lower level, aligned Key Results Areas (KRAs).
- 9 Four Deputy Commissioner Teams (DC Teams) have been established, each with its own portfolio of agencies. The intention is that a team will work with their agencies in facilitative ways to assist them develop the type and quality of practices that seems to be appropriate for the agency. In principle, the DC Teams reflect current views that central agencies should move from a command to an enabling model of operation.
- 10 Having said that, it is probable that, in the future, in Australia as well as New Zealand, budgeting and financial management will be expressed in output terms, while outcomes will be the focus of program and policy management.
- 11 There are also obvious parallels in this kind of thinking to emerging understandings of professionals-at-work, communities of practice and organisational learning: see, for example, Schön (1983), Senge (1990) and Wenger (1998).
- 12 Monitoring and evaluation should pay particular attention to the behavioural dynamics and effects of client use of the outputs delivered (for an extended argument regarding the importance of 'utilisation', see Ryan 1993).
- 13 The SSC is to be preferred to Treasury to avoid an economic and/or positivist skew to the prescriptions and practice. Cost-benefit analyses, cost-effectiveness analyses and the like are useful evaluative tools in some contexts, but the more important priority over coming years is to develop socio-political and policy approaches to evaluation.

References

- Barrett, P 1992, 'Evaluation as a strategic element of reform in the Australian public sector', in C O'Faircheallaigh & B Ryan (eds), *Program evaluation and performance Monitoring*, Macmillan, South Melbourne.
- Boston, J, Martin, J, Pallot, J & Walsh, P 1996, *Public management: the New Zealand model*, Oxford University Press, Auckland.
- Boston, J & Pallot J 1997, 'Linking strategy and performance: developments in the New Zealand public sector', *Journal of Policy Analysis and Management*, vol. 16, no. 3, pp. 382-404.
- CCMD (Canadian Centre for Management Development) 2001, *The learning journey: a guide to achieving excellence*, CMCD, Ottawa.
- Giddens, A 1984, *The social constitution of reality*, Polity Press, Cambridge.
- Giddens, A 1990, *The consequences of modernity*, Polity Press, Cambridge.
- Giddens, A 1998, *The third way*, Polity Press, Cambridge.
- Ministerial Advisory Group 2001, *Review of the centre*, State Services Commission, Wellington.
- Mintzberg, H 1994, *The rise and fall of strategic planning*, Free Press, New York.
- NZ Treasury 1987, *Government management: brief to the incoming government*, Wellington.
- O'Faircheallaigh, C & Ryan, B (eds) 1992, *Program evaluation and performance monitoring*, Macmillan, South Melbourne.
- Organisation for Economic Cooperation and Development (OECD) 2000, *Government of the future*, OECD, Paris.
- Organisation for Economic Cooperation and Development (OECD) 2001, *Government in the 21st century*, OECD, Paris.
- Robinson, M 2002, 'The public service and statements of intent', *Public Sector*, vol. 25, no. 1, March, pp. 6-9.
- Ryan, B 1992 'Evaluation in the Commonwealth Government: a critical appraisal', in C O'Faircheallaigh & B Ryan (eds), *Program evaluation and performance monitoring: an Australian perspective*, Macmillan, South Melbourne.
- Ryan, B 1993, 'Implementation and utilisation', unpublished manuscript.
- Ryan, B 1996, 'Are non-positivist approaches to evaluation replacing positivist ones?', paper presented to the AIC Conference, 'Performance-Driven Evaluation in the Public Sector', Sydney, 19-20 February.
- Ryan, B 1998, 'New tools for evaluators: process evaluation, program logic, intermediate outcomes and qualitative research methods', paper prepared for the Victorian Office of Premier and Cabinet, 22 December.
- Ryan, B 2002, *Managing for outcomes in the 21st century: a discussion paper*, Brisbane, Institute of Public Administration Australia, Queensland Division.
- Ryan, B 2003, 'Research and evaluation for the 21st century: from technocracy to learning', paper presented at the Ministry of Social Development Social Policy Research and Evaluation Conference, Wellington, 28-29 April.
- Schick, A 1996, *The spirit of reform: managing the New Zealand public sector in a time of change*, report for the State Services Commission and the Treasury, Wellington.
- Schön, D 1983, *The reflective practitioner: how professionals think in action*, Temple Smith, London.
- Senge, P 1990, *The fifth discipline: the art and practice of the learning organization*, Random House, New York.
- Shadish, W, Cook T & Leviton, L 1991, *Foundations of program evaluation: theories of Practice*, Sage, Newbury Park.
- State Services Commission 1998, *A better focus on outcomes through SRA Networks*, Occasional Paper no. 3, SSC, Wellington.
- State Services Commission 1998, *Looping the loop: evaluating outcomes and other risky feats*, Occasional Paper no. 7, SSC, Wellington.
- UK Government 1999, *Modernising government*, HMSO, London.
- Ussher, C & Kibblewhite, A 2001, *Outcomes focussed management in New Zealand: a background paper*, Treasury Working Paper, 01/05, Wellington.
- Weiss, C 1977, 'Research for policy's sake: the enlightenment function of social research', *Policy Analysis*, vol. 3, no. 4, pp. 531-545.
- Wenger, E 1998, 'Communities of practice: learning as a social system', *Systems Thinker*, June.
- Wholey, J & Hatry, H 1992, 'The case for performance monitoring', *Public Administration Review*, vol. 52, no. 6, pp. 604-610.



Insurance for AES Members



IC Frith & Associates can help you with your insurance requirements

- Exclusive Professional Indemnity Facility for Evaluation Professionals
- Workers Compensation Insurance
- Business Insurance
- Travel Insurance
- All other business insurance tailored to your needs

For an obligation free quote contact:

Brenda Hill
Phone: (02) 9899 2133; Fax: (02) 9634 2396
Email: brenda@icfrith.com.au

IC Frith & Associates Pty Ltd

ABN 48 097 088 287

Insurance Brokers