

# Development of program evaluation in Australia and the Australasian Evaluation Society – the early decades

## early decades

The year 2003 marks the 21st year of the series of National Evaluation Conferences (NECs) and international conferences, as well as the sixteenth year of the AES. This paper is not intended as a full history, nor as a dossier of the AES *per se*; that is for another time and place. Rather, it seeks both to reflect on whether there is a distinctive flavour of public sector program evaluation in the antipodes, and to give an overview of how evaluation developed in Australia, and to a lesser extent New Zealand. This will require the opening up and examination of the historical roots of program evaluation in Australia and of some aspects of current practice.

Colin A. Sharp



This paper, like the majority of the literature on which it is based, focuses on program evaluation in the public sector. There have been many other guises of evaluation in government, such as Corbett's (1991, p.3) 'industry of external review', and the long tradition of 'Inspectorates' in Australia's and New Zealand's education systems. I also had to draw the line between the public sector and the not-for-profit or third sector, although that sector has made important contributions to program evaluation in Australia, for example in grass roots self-evaluation.

### Precursors: program evaluation in Australia before the National Evaluation Conferences

Prescribed or compulsory government program evaluation in Australia is not a recent development. Indeed, there were a number of attempts in the 1950s and 1960s and earlier to introduce large-scale program evaluation as recommended or prescribed policy and practice, particularly in the community-oriented educational fields of Agricultural Extension and Technical and Further Education. Evaluation practices were also adopted in the social work and health disciplines by the early to mid-1970s. However, notwithstanding these and other early pioneers, program evaluation cannot be considered to have become a major government commitment or priority in Australia until the

watershed enquiries known as the Coombs Report and the Baume Report.

### The Coombs Report

The first of these reviews was the Royal Commission on Australian Government Administration (RCAGA, 1976), known as the Coombs Report or Coombs Commission<sup>1</sup>. Not only did this study pave the way for program evaluation, but it was also among the most instructive Australian government inquiries in identifying organisational diagnosis, and a form of benchmarking, as vital aspects of improvement of public sector administration. The Commission's Task Force on Efficiency described an agenda of reform, including performance audit and new public management, which took a generation to be addressed.

### Baume Report 'Through a glass darkly'

The other watershed in terms of evaluation came with the Fraser Government's commissioning of the Baume Report, the Report of the Senate Standing Committee on Social Welfare<sup>2</sup> (SSCSW, 1979a) entitled *Through a glass darkly*. Although it was a review of evaluation in Australian health and welfare services, it had far-reaching effects. Not only was this study an influential social reform vehicle, it also provided successive governments with a recommended definition of program evaluation for the whole of government, as 'the process of thoroughly and critically reviewing the *efficiency, effectiveness and appropriateness* of any program or group of programs' (SSCSW, 1979a, p.5, emphasis added).

In introducing the criterion of appropriateness, the Baume report suggested, *inter alia*, that appropriateness was 'the extent to which the program is in accordance with the moral, ethical and social norms of the time...' (Gross, 1979, p. 58, italics original). However, with governments becoming more wary of the independence and authority of the public sector, the interpretation of appropriateness soon became firmly grounded in 'Government policy' (Department of Finance, 1992a).

As well as these Commonwealth Parliamentary inquiries, there were a number of state government and commonwealth departmental and non-governmental reviews. The Australian Schools Commission established a national system of research and evaluation for its new programs (see below). The move to devolve autonomy and management to schools was accompanied by institutionalised self-evaluation as well as building a higher profile for evaluation in government policy and administration (see Caldwell, 1998; 2002; Wall, 1987). Similar developments in school-based evaluation occurred in Tasmania (Caldwell, 2002). The government of Tasmania also commissioned a review of evaluation methods which could be applied in its welfare services (see Jamrozik & Hooke, 1975).

At the departmental level in the Commonwealth Government, the Department of Employment and

Industrial Relations (1977) promoted a debate on the underlying philosophies of evaluation with its report *Evaluating the Helping Services*. In the same year the Australian Council of Social Service established an Evaluation and Accountability Task Force to facilitate a capability and commitment to evaluation among the local level non-government agencies. This had a profound influence on the Senate Standing Committee inquiry, especially in the establishment of the definition of evaluation used in the Baume Report, which subsequently influenced the Commonwealth Department of Finance's concept of evaluation for the whole of government.

In tracing the history of evaluation in Australia from Federation to the Whitlam Government, the Baume Report rather overstated the impression that it 'showed that there was almost no formal evaluation during this time' (SSCSW, 1979a, p. 19). Nevertheless it was a positive contribution to articulate that the main prerequisites for systematic planning and evaluation are:

- determination of needs;
- delineation of goals and objectives for programs and groups of programs;
- delineation of criteria or standards for evaluating progress toward those objectives and for assessing the competing claims of proposed programs;
- development of a database for providing measures of those criteria, through a process of monitoring the program; and
- application or appropriate use of findings.

The report further noted that when social welfare programs were formulated and implemented before 1970, 'one or two of these prerequisites were observed sometimes but seldom all five together' (SSCSW, 1979a, p.19).

The Baume Report also commented on the 1969 Nimmo Report (the Commonwealth Enquiry into Health Insurance) as being the 'best evaluation document' up to 1973. While the Nimmo Report met at least some of their criteria, it was 'far from ideal as an evaluation report. However, considering that evaluation was not widely discussed and understood at that time, the inquiry was certainly a step forward.' (SSCSW, 1979a, p. 22).

However, the Baume review did not mention the Royal Commission on Public Service Administration (1918-1920) or the Royal Commission to Consider and Report Upon the Public Expenditure of the Commonwealth of Australia (1918- 1921), although both met at least a few of the criteria noted above. Similarly, the report overlooked a considerable body of evaluations in agricultural extension and TAFE, and the importance of the Schools Commission in establishing a national educational evaluation system.

The key claim in the report which I wish to reconsider here, that 'there is no tradition of evaluation in Australia' (SSCSW, 1979a, p.109), may have been influenced by comparison with the

USA and by the focus on the health and welfare sectors. Nevertheless, during the period of the Baume Report there seemed to be an increased interest in the establishment of evaluation within government organisations, including the Commonwealth Department of Social Security, the South Australian Department of Community Welfare and the Western Australian Department for Community Welfare, as well as within related professions such as social work.

Although all these evaluation-related reviews, and subsequent development of governments' evaluation capacity, served to lift the level of commitment to, and the quality of, program evaluation, they still were piecemeal and demonstrated the point made in the Baume Report about the lack of a consistent framework and culture of evaluation. Subsequent to the Baume Report, there were various attempts to develop a consistent performance evaluation and management improvement framework in some state and national programs in Australia, but it was not made mandatory or endorsed as a whole-of-government policy until the late 1980s.

Although all these evaluation-related reviews... served to lift the level of commitment to, and the quality of, program evaluation, they still were piecemeal and demonstrated the point made in the Baume Report about the lack of a consistent framework and culture of evaluation.

In three sub-fields of education, however, program (formative) evaluation became the prescribed (eventually mandatory) and widely implemented policy of the government agencies concerned. Therefore these fields (agricultural extension studies, state primary schools and TAFE) invite further attention regarding the question of distinctiveness, and perhaps to discover some lessons for today.

### The era of Agricultural Extension Studies

Like the role of agricultural research in the development of statistical methods, agricultural extension provided opportunities which had a formative effect on the early approaches to program evaluation. Some of the current practitioners of program evaluation in the human services gained their experience through agricultural extension program evaluation. For example, Michael Quinn Patton was involved for many years in extension services through the Institute of Agriculture, Forestry and Home Economics at the University of Minnesota and was the Editor of the *Journal of Extension*.

In the USA, it had been the 'dust bowl' storms of the Great Depression that had laid these foundations (see Farquhar, 1961; 1963; Trotman, 2003). In Australia the Murrumbidgee Irrigation Area (MIA) Agricultural Extension was formed by the CSIRO, State and local authorities in 1947, with a deliberate

research and evaluation strategy to test this program (Farquhar, 1961). The evaluation, carried out by designated evaluators, was reported in 1952 by the Bureau of Agricultural Economics, and 'many lessons were learned from the careful evaluation of the MIA extension experiments ... probably the most significant was the lesson in agricultural organization' (Farquhar, 1961, p. 210).

In the 1950s two social psychologists from the University of Melbourne, Professor Oscar Oeser and Dr Fred Emery, were active in sociological and psychological studies of rural communities as part of a UNESCO world comparative study (Oeser & Emery, 1954). As a result of this association with CSIRO and the University of Melbourne Faculty of Agriculture, Oeser and Emery (1958) conducted a landmark study, published as *Information Decision and Action*, on the diffusion of innovation and decision-making among farmers.

In 1962 the CSIRO (on behalf of the Australian Agricultural Council) organised the second Federal Conference on Agricultural Extension Services. Attendees were interested in, and drew from, the tradition of agricultural extension in the USA where some of the early roots of program evaluation lie. Indeed, one member of the conference organising committee, Dr. Reginald Farquhar saw the benefit of grass-roots formative evaluation in agricultural extension programs in the USA, and advocated a similar approach in Australia.

In the heyday of the rural sociological studies, the University of Melbourne *Notes on Agricultural Extension* began in 1966 and formed the basis of some of the early debates and practices which were the forerunners of more formalised program evaluation. A useful summary of the early years of this work is provided by Parkin (1972).

The use of a standard model of program planning, specifically including evaluation, was introduced as early as 1956 in Queensland, and training in this model and its implementation became compulsory in NSW in 1968. Parkin elaborated on these processes by identifying the role of evaluation in linking the various stages, as the following 'Guidelines for Planning of Programs' showed, in four steps:

- **Situation Analysis**
  - Long term – fixed component
  - Long term – variable component
  - Short term – variable component
- **Objectives**
  - Long term
  - Short term
- **Workplan**
  - Objectives plan
  - Calendar of operations
- **Evaluation.**

Two key issues of the Agricultural Extension era which may provide clarification on current concerns in program evaluation were the prescription of formalised evaluation, and its professionalisation.

The prescription of formalised program evaluation in agricultural extension

It is not clear whether there ever really was a government directive to formalise program evaluation in most agricultural extension services. It seems that the strong conformity arose by virtue of the nature of the project work funded by the Commonwealth grants, which made extensive use of questionnaire survey feedback methodology, and the nationwide normative use of a uniform program planning and management model (see Parkin, 1972). However, evaluation was recognised to have an important role early in the development of the field.

### **The professionalisation and status of agricultural extension**

By the 1962 national conference there were calls for professionalisation of agricultural extension. Indeed, part of the rationale and aims of that conference was 'recognition of the need to develop further the profession of extension, to define qualities needed in extension workers, and to train extension workers accordingly'. However, there are some who believe that agricultural extension was always borderline as a profession, and ambivalent about its professional role, as illustrated by the myth that the extension expert was taught to 'work yourself out of a job' (Cary, personal communication, 1992).

Eventually the 'profession' did not survive the test. When the specific commonwealth funding was stopped and redistributed under the 'New Federalism' of the Fraser Government, the State governments spent the former extension allocations elsewhere.

Woog (1990) argued that agricultural extension went through a major change from the late 1970s. He claimed that 'improvement did not occur for decades in agricultural extension practice because evaluation in that field emphasised methods and techniques without putting them in context. It has been further argued that where human activities are concerned, context must include the subjectivity and transient nature of human systems and the diversity of human values' (p.247). In effect, he was emphasising that agricultural extension officers needed to develop skills in self-evaluation, 'critiquing' or 'meta-evaluation'.

### **Evaluation in Education**

More than ten federal and state-sponsored reviews between 1971 and 1984 inspired Byrne, Houston and Thomson (1984) to write that 'government inquiries into education in Australia have been one of the few growth industries of recent years.' Kerith Wall's research for a Master in Education gives a local insight into the waxing and waning of program evaluation in the education system over a hundred years from the 1880s.

According to Wall's review, an issue from the 1890s to 1920s was the 'payment by results' scheme which brought a performance appraisal function to the program evaluation role of school inspectors. The major change in program evaluation came with the Ramsay Report of 1960, which recommended

that inspectors continue their inspectorial duties, but become more like advisors on professional matters (such as program evaluation) rather than assessors for performance appraisal of individual teachers.

The 1973 Karmel Report led to the establishment of the Schools Commission, which recommended seven programs in education across the country. From this came a national system of programs which were systematically researched and evaluated under the surveillance of a National Evaluation Committee, which had a significant budget and explored the wider implications of innovations as well as evaluating their effectiveness. This institutionalised evaluation in government policy and administration as never before in Australia.

The 1980s saw a trend towards devolution of school management, which was accompanied by an emphasis on self-evaluation of school curriculum and other programs. In 1983 the newly elected Cain Government made program evaluation in Victorian state schools the primary responsibility of the School Council, as a part of a process of devolution of accountability towards the community (see Caulley, 1988, Toner, 1983). Indeed, the Cain Government signalled that this devolution also meant a drive for efficiency and effectiveness using program evaluation in education. Other social services, such as community welfare services (see Toner, 1983), were soon to follow suit.

TAFE was one of the first mainstream public sector functions to make a strong commitment to organisational and system-wide evaluation, including prescribing program evaluation as a national system. This was separate from curriculum evaluation as part of course development, etc., which was also well established in TAFE.

The mid-1970s saw several reviews and policies which began to signal evaluation as a key process for consolidating TAFE and as a means of legitimising its accountability and effectiveness. In particular, the South Australian Department of TAFE sponsored a review of schools and colleges which led to a national pilot study to introduce a uniform approach to evaluation of TAFE institutions across the country. Again the approach was based on the premise of increasing accountability and value for money, but couched the presentation in terms of 'persuasion' rather than 'model' in recognition of staff suspicion of evaluation. There were commonly references to the role of the 'Evaluator' and the needs for objectivity, impartiality and accountability, and for staff to be separately identified as 'Evaluators' (Byrne et al, 1984a; Kuhl, 1978).

In general terms it is instructive to consider Kemmis' reflections on this era of evaluation in education, as he was not only a participant but a critic. His experience led him to realise that the traditional (1970s)

belief that program evaluation is a specialised technological process ... has weakened the obligation and commitment of program participants and decision makers to critical self

reflection, and has created conditions in which specialist evaluators (rather than program participants themselves) now carry primary responsibility for determining the efficacy and effectiveness of programs. The ascension [sic] of evaluation as a specialist professional field ... has been accompanied by a decline in the commitment to developing more powerful forms of self-evaluation. (Kemmis, 1989, pp. 3-4)

## The era of the National Evaluation Conferences

### The early National Evaluation Conferences

Although there had been other conferences and organisations dealing in part with program evaluation, it was Dr Anona Armstrong's vision and leadership which brought the disparate groups together in the first of the (initially biennial) National Evaluation Conferences (NEC) in August 1982. At that time the newly elected Cain Labor Government had started to review services, and Dr Armstrong and Professor Alex Wearing were advising the Victorian Department of Community Welfare Services on setting up their social indicators for program planning and evaluation.

The 1982 conference was intended to 'review the "state of the art" in Australia' (Armstrong, 1983, p. i). The fourteen speakers presented papers on evaluation applications in health, education, welfare, and manpower planning. A session was also conducted on 'managing evaluation', and a panel discussion summarised the key issues in closing. There were 93 participants from all the States of Australia.

Two of the papers provided a historical perspective. Bruen considered the history of health care evaluation in Australia to be:

remarkably similar to the history of Australia's economic policy. No need for any evaluation in the steady laissez-faire days of Menzies. Indication that all was not well in the troubled time of Holt, Gorton and McMahon, rapid expansion in the heady days of Whitlam, contraction and even recession in the difficult days of Fraser and the razor gang. Throughout this period, evaluation has played an influential, though not always an obvious, role in the health decision-making process. Perhaps it is not always easy to point to specific results of individual projects, but the cumulative effect of health services research carried out over the last twenty years has, I believe, been considerable. (Bruen, 1983, p. 49)

However, he identified problems with the kind of health care evaluations conducted so far:

There has, over the last ten years, been some realisation in the health field that not all evaluation is worthless unless it conforms to the rigours of

experimental design. Unfortunately there has not always been a willingness to use the most rigorous, and therefore the most informative type of evaluation possible in any specific situation.

We now recognise that formative evaluation (i.e. evaluation of the functioning of, for example, a community health centre) is a needed and useful activity, without the evaluation having to shed light on whether or not community health centres as a whole are a good thing or bad thing for Australia. At the same time, evaluation studies of major health areas (e.g., hospitals, Family Medicine Program, medical research) have often turned out to be nothing more than a collection of opinions expressed by the authors, illustrated with a few selectively chosen facts and figures. Often the time taken by these studies and their cost are more than would be incurred by a more rigorous attempt to evaluate the subject at hand. The Commonwealth, I'm sad to say, is as much at fault as anyone in this regard. (p. 53)

Ralph Straton noted that, notwithstanding the high level of interest in the role of evaluation as a means for achieving efficient and effective services, the level of funding had remained small and was directed substantially toward small local studies, often conducted by practitioners. While expressing optimism about the benefits which could flow from evaluation and the interchange of ideas among practitioners from diverse disciplines, Straton noted that there had so far been a limited sharing of ideas which might contribute to the development of practice.

The 1984 National Evaluation Conference gave opportunities for more comprehensive debate about the direction of the role of evaluation in government policy which had been raised by the Baume Report and by Victoria's stance on Program Budgeting. Indeed Mune (1986) and others were quite critical of the emphasis in the Baume Report on accountability of programs to the Parliament for efficient use of resources.

### The rise of managerialism

Mune's views seem ironic today, as the now dominant 'managerialism' regime had yet to take hold in Baume's period. It was not until the Reid Report of 1983 promoted management by objectives and performance control, leading to the Financial Management Improvement Program (FMIP) in 1984, that the full course of managerialism was set in train in the Commonwealth. It was this program that fostered an 'Evaluation Strategy' in 1988 and mandated annual program evaluation plans and reports for all Commonwealth-funded programs.

In effect, managerialism meant devolving responsibility from the central agencies down towards the service delivery interface, and 'letting the managers manage', as well as lifting the importance of program evaluation and internal audit with tighter budgetary accountability, i.e. 'making the managers manage' (e.g. Keating, 1989).

In general, Australia seemed to be following a trend of several of the OECD countries. According to the Department of Finance's 1988 FMIP Report, the United Kingdom instituted a similar system in 1982, as did Canada in 1986 and New Zealand in 1989. Of course, the Johnson Administration in the U.S.A. in 1965 promulgated what was probably the original form of Program Budgeting, derived from the Defence Department's Program Planning and Budgeting System (PPBS), which gave rise to their strong emphasis on program evaluation.

These changes were vital ingredients in the making of the National Evaluation Conferences. Some sections of the public sector contributed by strengthening the evaluation capability of the central agencies (e.g. the Commonwealth Public Service Board built an Evaluation Unit of over a dozen staff), and some by developing their own in-house capability. For example, in April 1983 the Commonwealth Rehabilitation Service (CRS) created a position called Research and Evaluation Manager (the first incumbent was the author). But it did not seem appropriate for a small human service organisation to build an empire around evaluation. So I went about developing quality circles and internal capability among the CRS staff for self-evaluation of the programs developed by allied health staff. This led to the creation of a consultancy for Jerome Winston to develop a self-evaluation program using examples from the CRS and an 'Evaluation Training Kit' (Sharp, Winston and Bhagwandas, 1986).

### The foundation of the AES

The 1986 NEC in Sydney became the turning point with the formation of the Australasian Evaluation Society (see Armstrong, 1987). The Canberra National Evaluation Conference of 1987 saw the more personal side of the debate about the emphasis on appropriateness, access and equity as criteria in the evaluation of public services, versus the advocacy of efficiency and cost-effectiveness. These sides were politically drawn between the Public Service Board, in the person of its Head Dr Peter Wilenski, and the Department of Finance (DoF), under Dr Michael Keating.

It was more than ironic that Dr Wilenski was the invited opening speaker of the Conference, in the final week before the demise of the Public Service Board and his transfer to head the Department of Transport and Communications. This was part of the Hawke Government's 'Bastille Day' restructuring of departments (see Pusey, 1991) which saw the closure of the program evaluation unit of the Public Service Board and the transfer of some of its staff to the Department of Finance. As Dr Wilenski put it (during several ad-libs), formerly 'evaluation belonged to the Public Service Board' which pursued management improvement, now 'we've lost the battle to the Department of Finance' which had taken evaluation over for its purposes of 'cost cutting' (Wilenski, 1987).

### Government evaluation after the Baume Report

The decade after the release of the Baume report saw a number of studies and reviews of current government evaluation practice which gave widespread coverage to the need to document and review public sector reforms. These included:

- *Not Dollars Alone: Review of the Financial Management Improvement Program* (Australian Parliament, 1990). This report paved the way for recognising the importance of non-financial performance measurement;
- *The Australian Public Service reformed: An evaluation of a decade of management reform* (MAB-MIAC 1993). This study reviewed the whole of the Commonwealth Government's managerialist reform process;
- *A Model for Best Practice in Client Service in the Public Sector* (Joint Council of Australian Public Service, 1993) put service quality high on the agenda.
- *Accountability in the Public Sector* (MAB-MIAC 1993) established a uniform model of accountability for the whole of Government;
- *Beyond Bean Counting: Effective Financial Management in the APS - 1998 & Beyond* (MAB, 1997) consolidated the importance of non-financial data in both monitoring and evaluation of Government programs, and also promoted the use of a Balanced Scorecard.

Resource allocations for evaluation activity varied from delegation, to the discretion of the Chief Executive Officer of the Department (e.g. in South Australia), to specific budgetary requirements.

According to the 1994 ANAO efficiency audit of evaluation in Commonwealth programs, some Commonwealth Government portfolios

spent as much as 2% of their multi-million dollar budgets on evaluation of their programs. Most Governments had delegated the oversight of their evaluation strategy to a particular central agency, e.g. DoF for the Commonwealth; or else, as in most of the States, there were evaluation units established in large departments or agencies, although they may have been called something like 'Quality Assurance Unit', as in the NSW Department of School Education.

The 'lead agency' role in the Commonwealth in the 1980s led to the imposition of a strong set of espoused values from the top.

### 'Evaluative Culture'

The 'lead agency' role in the Commonwealth in the 1980s led to the imposition of a strong set of espoused values from the top. For example, the Secretary of DoF asserted (Sedgwick, 1994, p. 20) that the Australian Public Service (APS) was becoming an 'evaluative culture', which he described as an 'attitude, body of technology and incentives'

that led to a questioning by managers and ministers of the effectiveness, efficiency, and appropriateness of government programs.

However, *The Australian Public Service Reformed*, while recording a high degree of satisfaction with the general thrust of the progress in management improvement, expressed concern about the so-called 'evaluative culture' that Sedgwick espoused:

The evaluation strategy has been operationalised into a set of procedures that agencies have grown accustomed to, and they are undertaking a great deal of programme evaluation activity. An irony here is that, once again, there may have been a tendency to focus on the procedures, not the results. Everyone is doing evaluation, following the procedures laid down. The success of the process is being measured in process terms... not in terms of program outcomes. ... In sum, the reforms are only half-complete. ... many of the changes already put in place are still taking effect and, even with no further direct initiatives, will lead to further improvements in agency performance in due course. (MAB-MIAC, 1993a, pp. 472– 473)

Although this may have detracted from Sedgwick's organisational diagnosis, it actually gives some credence to the suggestion that Australia (at least in the Commonwealth public sector) had developed some form of whole-of-government evaluation strategy integrated into management improvement.

### International Comparisons

By 1984, the Office of Management and Budget (OMB) of America's Executive arm of government had rescinded its Circular A-117 *Management Improvement and the Use of Evaluation in the Executive Branch* (Wye, 1992). This was just at the time when in Australia, the Commonwealth and most of the States had begun to establish program evaluation strategies as a tool of management improvement. Within the next ten years, all government agencies (Commonwealth, State and Territory) were basically required to produce evaluation plans and evaluation reports based on regular performance monitoring of programs in the budget cycle, as well as specific evaluations of any major initiatives. This earned Australia a reputation in OECD circles, and especially in the USA, for integrating performance information, with a focus on outcomes of services for clients, into the budgeting and management cycles of government. This perception has been of interest to the American OMB and National Performance Review (NPR), and in shaping their Government Performance and Results Act (GPRA) of 1993, according to one of the senior government advisers<sup>3</sup>, who wrote that:

'OECD studies reported that Australia was some five to ten years ahead of the United States' experience of performance measures, and that

their centrality to the entire process of reform has greatly accelerated their moves towards achieving a real environment for 'managing for results.' ... In testimony before the Senate Governmental Affairs Committee in May 5, 1992, OMB's Deputy Director for Management judged Australia to be 'at the leading edge in performance measurement' (Breull, 1994, p.4).

However, it is important to recognise that some of what appears innovative was an adaptation by Commonwealth agencies of various approaches from the USA (e.g. linking evaluation to Program Budgeting) and the Australian States, some of which had faded away in response to changes of government. For example, South Australia under the Tonkin Liberal Government was probably the first State to introduce a form of program budgeting in 1982, but it was never fully pursued as a performance management system in the public service, partly due to change of government in 1985, and by 1990 there were significant performance management and program evaluation strategies which superseded the Program budgeting approach.

### Has an evaluation profession emerged in Australia?

It is not intended here to evaluate the AES or those who call themselves 'Evaluators' according to some set of criteria of "profession". However, it is relevant to mention the early concerns for professionalisation which are to be found in the various communications of the NECs and the AES.

In his keynote address to the National Evaluation Conference in 1989, Michael Quinn Patton, then AEA President, pointed out that the AEA advocated promoting evaluation as a profession. He also suggested that for evaluation to thrive as a profession in this country and worldwide, we needed 'vision, quality products and processes, and skilled, trained evaluators.' He also said:

My remarks are meant to be particularly targeted at the members and leadership of the Australasian Evaluation Society because the Society has a special role to play in the future development of the profession. (Patton, 1989, p. 40)

There were some (such as Jerome Winston and I) who took issue with the use of the term 'Evaluator' in the Australian context and were concerned about the restrictive or elitist connotations of 'professionalising' evaluation. This view was reflected in the 1995 Draft Strategic Plan of the AES and in a resolution from the 1994 Strategic Planning Workshop to the effect that the AES should not be an elitist exclusive guild of professional 'Big-E' Evaluators; rather its role should be to encourage those who are interested in the theory, practice and use of evaluation.

In the USA, the spread of program budgeting to all government departments in 1965 under the Administration of President Johnson created an

industry of in-house experts, policy analysts and program evaluators in the public sector. It has been estimated that in about five years there were at least 800 such analysts in 16 domestic agencies (Marvin & Rouse, 1970). But despite the impression that might have been conveyed by the Anona Armstrong's publication in 1983 of a *Directory of Australian Evaluators*, Australia did not at that time have a tradition of professional Evaluators in either the public or the private sector.

The role of the private sector has been difficult to identify as a force in the AES. There have been numerous private consultancies or companies conducting evaluations, but no coherent group of private sector interests (consultants or commerce/industry) or any sense of a distinctive private-sector approach to evaluation in this country. This contrasts markedly with the important (even dominant) role of the private sector providers of evaluation consultancy in the USA. Although, it has been suggested that there is an Australian 'brand' of management consultancy, there is little to indicate a distinctive style in evaluation consultancy.

An example of ambivalence about professionalism is the periodic revival of self-evaluation (by managers and client interface supervisory staff) as a process of making the role of the 'Evaluator' redundant. For instance, in SA the evaluation strategy introduced by the Premier in May 1990, with mandatory reporting of evaluation for 50% of programs over five years, was specifically intended not to create an industry of Evaluators. Indeed, this strategy was supported by a training program designed to give the public sector managers an introduction to conducting their own evaluations.

The aspects of professionalisation on which a consensus is developing in the AES are ethics and standards in evaluation practice, and skills development through adequate training. While most current members will be aware of the developments in both fields over the last few years, it may be of interest to look briefly at their origins in the late 1980s and before.

## Ethics and Standards

The first AES Working Group on Ethics and Standards was formed in October 1988 in response to a complaint about an evaluation being conducted by consultants (not members of the AES). Earlier that year Jerome Winston became aware of the complaint and raised the issue with Bryan Lennie, the President at that time. They convened a group comprising Jerome, Professor Stephen Kemmis and me to investigate and consider the AES response on this issue. The Working Group gathered relevant information and made representations individually to the various interested parties, and the issue dissolved without direct AES action, though not necessarily satisfactorily so far as either the complainant or the AES was concerned. Nevertheless, it was decided to take up the broader concern over the role of the AES and to address the need for some kind of code of conduct for evaluation.

With the approval of the AES Committee, Jerome Winston convened a plenary Forum on Ethics and Standards at the National Evaluation Conference in 1989, which I chaired. The Committee on Ethics And Standards in Evaluation was formed to review relevant codes of practice and develop some position papers, which were presented at the National Evaluation Conference 1991 in Adelaide. In July 1992, an Interim Code Of Ethics was circulated to AES members and further developed by Chris Milne, Ian Trotman and others into the *Guidelines for the Ethical Conduct of Evaluations* (AES, 1996).

## Training and Professional Development in evaluation

According to the Baume Report, there was some recognition at the time of the need to expose senior public servants to evaluation activity and training by way of a seminar series offered to public servants on an Executive Exchange scheme through the Canberra College of Advanced Education (1978– 79). Similar exchanges occasionally involved outside academics: for example, Jerome Winston at Phillip Institute of Technology used such a scheme to work in the Public Service Board and then the DoF in Canberra from July to December 1987 (arriving just at the time when the Public Service Board was being abolished).

By the start of the 90s, some Commonwealth Departments were conducting their own training programs in evaluation methods (e.g. the Commonwealth Rehabilitation Service). A number of the States' central agencies ran introductory training programs in evaluation – NSW through the Public Service Board's Program Evaluation Unit, South Australia through the Office of Cabinet and Government Management.

By 1990, when Jerome Winston edited a *Directory of Evaluation Training Opportunities* in Victoria, there were already at least three graduate diplomas in program evaluation, including offerings at Phillip Institute, Latrobe University and the University of Melbourne. Dr Armstrong's company ET&S also offered a workshop program.

As a recognition of the need for training and development at that time the AES (proposed by President, Dr Caulley) created the positions and subcommittees of a Training & Professional Development Committee, initially co-ordinated by Associate Professor John Owen at The University of Melbourne.

## Is there a distinctively Australian Approach to program evaluation?

Several authors have examined whether Australia and New Zealand have developed distinct styles of evaluation practice:

- Darrel Caulley (1992), when he was AES President, suggested that there had been five generations of (program) evaluation, and that the fifth generation (derived from action research) was emerging as a distinctive form of

'self-evaluation', which had been encouraged within the Australian public sector.

- Love (1995) explored the similarities and differences between Australian/NZ and Canadian evaluation practices, reinforcing the view of other OECD commentators on the progressive approach that Australian and New Zealand Governments have taken to evaluation.
- A less flattering comparative analysis was recently provided by Ryan (2003) who recommended that New Zealand should avoid the 'death by evaluation' which had resulted from the Australian Commonwealth Evaluation Strategy of 1988.
- Rogers (2000, p.1) identified four features of Australasian evaluation practice – eclecticism in evaluation methods and practice, encompassing internal evaluations as organisational capacity building, performance indicators as a system requirement for programs, and an effort to build evaluation into programs for, and by, indigenous people.

Based on the experience of reviewing the evaluation literature for my earlier (tenth anniversary) 'Interim History' project, my impression is that there have been instances over four decades of innovative combinations of elements of grass-roots self-evaluation for bottom-up program planning and improvement, and elements of mandatory top-down accountability and reporting. Whether this is sufficient to confirm a distinctive Australian (or Australasian) style of evaluation is debatable.

Many commentators agree that there are more similarities than differences between Australian performance management systems and evaluation practices and those which apply in the US, Canada or Europe. Many of the differences, moreover, are a consequence of the particular local Australian cultural and politico-legal contexts for evaluation. Fundamental factors like the absence of a Bill of Rights, the heritage of British Common Law, the Westminster Parliamentary tradition and the historical predominance of primary industry in the Australian economy will be well known to readers, so it is appropriate here to concentrate on some of the other distinguishing factors and highlight some of the more important developments:

- The introduction in 1984 of the Commonwealth FMIP and Evaluation Strategy. FMIP required outcomes-oriented program budgeting and evaluation, focused on the evaluative criteria of efficiency, effectiveness and appropriateness of a program or service, and promoted a consistent evaluation framework for more than a decade.
- A central agency-driven emphasis on internal formative evaluation, with the integration of accountability and performance improvement into the responsibilities of managers. This mandated evaluation in all Commonwealth-funded programs, requiring annual Portfolio

Evaluation Plans (from 1991), Program Performance Statements, and periodic efficiency audits by the Australian National Audit Office of evaluation practice in the Commonwealth.

- A Government-sponsored<sup>4</sup> forum for promoting evaluation theory, research and practice in the guise of the National Evaluation Conferences, subsumed since 1986 by the annual AES conference. This forum has facilitated cross-fertilisation of evaluation practice throughout Australia and New Zealand, and the formation of a diverse practitioner base, while rejecting the opportunity of becoming a guild of professional evaluators.
- A commitment since 1988 at the highest levels of Government and the public sector (Commonwealth and most States), and in much of the private sector, to quality improvement and international standards of accreditation on quality management such as ISO9001.
- Agreement through the Council of Australian Governments (COAG) to promote a common framework for performance indicators and benchmarking for Government Business Enterprises (e.g., gas, electricity, water utilities) and the major human services (e.g., schools, correctional services, health).
- A national review to promote quality management, performance improvement and accountability in the third sector.

## Conclusions

It is appropriate to end this review with the words of an Australian leader in program evaluation, Professor Stephen Kemmis, summing up the 1986 NEC:

'We have begun to see a collision between views of evaluation as... a process of gathering information for bureaucratic-administrative decision making, on the one hand, and evaluation for a more generally educative purpose, on the other. Some here are interested only in the one view, and some only in the other; some believe that the two can be reconciled in a more liberal eclectic and pragmatic view of evaluation, and some that the two are permanently and essentially in opposition. Those who take the oppositional view see evaluation for bureaucratic-administrative decision making and evaluation for participant education as contrary poles in an opposition between conflicting interests in the modern state...' (Kemmis, 1989, p.2)

Since then, the 'oppositional view' seems to have gained the upper hand in the process of governments making program evaluation mandatory, and as a result there is a 'very general tendency in our evaluation methods and studies towards service to the state alone' (Kemmis, 1989, p.3).

It is relevant to consider the lessons of the Agricultural Extension era (especially the 1960s) when a program planning model, including

evaluation, became uniform policy across the States under Commonwealth sponsorship. From this experience the AES might be concerned about:

- the covert culture of evaluation as a self-improvement tool kit for managers in Australia, with fragmentary roots dating before the current 20-year era of program evaluation;
- the experience that like other evaluation models, this one was transitory, and vulnerable to changes in government policy or major structural changes; and
- the political issue of professionalisation of evaluation.

As Sue Funnell (1989, p. 43) pointed out in her review of the 1989 AES Conference, Patton had 'thrown down the gauntlet to our association with some challenges for the years ahead'. Indeed, Patton left the AES with the challenge of creating a view of evaluation as a 'profession', but he did not adequately address all the implications of eliteness and control, and the necessities of ethics and standards. However, in pursuing that vision, members of the AES might be wise to reflect on the past and consider other perspectives than the current regime.

The challenges include the following:

- Can the AES provide a balanced view and a 'marketplace of ideas' which fairly serves the various stakeholders?
- Under the economic-rationalist paradigm, and in these days of 'leaner government', when 'we can serve our political and bureaucratic masters by finding out where programs can be cut and savings made' (Kemmis, 1989, p. 3), can members avoid 'Shylock's problem ... the point at which we cannot get the pound of flesh without destroying the program itself' (Kemmis, 1989, p.3)?

These issues are appropriate for debate in the AES Conferences and in the *Evaluation Journal of Australasia*, but the Society needs a collection of sources which the debating parties may draw upon for their information. Without an adequate historical perspective these debates will continue to emerge and fade, as individuals come and go and corporate memory fades.

## References

- Caldwell, B. J. 1998 "Management of Schools: Devolution of Responsibility." Paper delivered at ACER Conference, [online [http://www.edfac.unimelb.edu.au/EPM/papers/BCaldwell/ACER\\_Paper.html](http://www.edfac.unimelb.edu.au/EPM/papers/BCaldwell/ACER_Paper.html) accessed 15/10/03]
- Caldwell, B. J. 2002 "Self-Managing Schools in Australia: Outcomes and Outlook." Paper delivered at ACSSO Conference, [online [http://www.edfac.unimelb.edu.au/dean/docs/ACSSO\\_paper.pdf](http://www.edfac.unimelb.edu.au/dean/docs/ACSSO_paper.pdf) accessed 15/10/03]
- Corbett, D.C., 'Contours since Coombs- The changing landscape in program evaluation' in J. Uhr (Ed.) Program Evaluation. Canberra: The Australian National University, Federalism Research Centre, pp. 3-18.

- Department of Employment & Industrial Relations, 1977 *Evaluating the Helping Services*. Canberra, Australian Government Publishing Service.
- Department of Finance (DOF) 1992a, *Evaluation Papers: Introduction to the Series and Overview*. Evaluation and Statistical Services Branch, Canberra.
- Farquhar, R.N. (1961). "Extension Education in Australia: Extract from "Comparative Extension Education in Anglo-Saxon Countries with Particular Reference to Australia". New York, Cornell: 170-217.
- Gross, P.F., 1979, "Recent evaluation activity in Australia". Commissioned paper by the Senate Standing Committee on Social Welfare, (SSCSW) 1979b *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 2 Papers Commissioned by the Committee*. Canberra, Australian Government Publishing Service, pp. 41-59.
- Jamrozik, A. & Hooke, R., 1975 *Report on the Evaluation for the State of Tasmania*. Hobart: Tasmanian C.A.E.
- Parkin, R.J., 1972, 'An approach to program planning in agriculture extension in Australia.' *Melbourne Notes on Agricultural Extension*. July, no. 7, pp. 5-97.
- Senate Standing Committee on Social Welfare, 1979a *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 1 The Report*. Canberra, Australian Government Publishing Service.
- Senate Standing Committee on Social Welfare, 1979b *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 2 Papers Commissioned by the Committee*. Canberra, Australian Government Publishing Service.
- Trotman, I. 2003, "Evaluation in New Zealand: A founder's reflection". In Lunt, N., Davidson, C. & McKegg, K. (Eds.) 2003. *Evaluating Policy and Practice - A New Zealand Reader*. Auckland, Pearson, pp. 21-26.
- Wall, K. 1987 *Total programme evaluation in the primary school*. Unpublished M.Ed. Thesis, LaTrobe University Melbourne.

## Endnotes

- 1 Chaired by Dr H. C. Coombs, this Commission began during the Whitlam government but was completed under the Fraser government.
- 2 This committee included probably the two most powerful reformers of health (both were medical practitioners) and social welfare of the subsequent decade, Senator Peter Baume (Liberal NSW), and Senator Don Grimes (Labor, Tas.) (cf. Baume, 1977).
- 3 Arnold Love (1995) corroborated this story.
- 4 The Commonwealth and all State Governments in Australia and individual agencies in New Zealand (e.g. the Education Review Office) have sponsored AES and National Evaluation Conferences, most notably by in-kind support with the secondment of staff as conference coordinators. This trend escalated with bigger budgets and more staff resources, peaking in Victoria in 1992. Conversely, when evaluation faded as a priority of the Commonwealth and some States in the late 1990s, this decline was reflected in a cessation or reduction of sponsorships. For example, although the 1991 Adelaide Conference got 1.5 staff and a significant budgetary subsidy, the 1997 Adelaide conference got nothing from the SA Government. However, the following year in Wellington there was a major sponsorship in money and staffing.

## Chronology of key events 1918-1991

- 1918-1921 Royal Commission to Consider and Report Upon the Public Expenditure of the Commonwealth of Australia
- 1956 Concepts of Program Planning introduced in Queensland; other States follow from 1957 through to 1965
- 1969 Commonwealth Committee of Enquiry into Health Insurance (Nimmo Report)
- 1971 Australian Association for Research in Education (AARE) held its first annual national conference with guest speakers from the USA on research and evaluation methods, debating the merits of qualitative vs quantitative methods.  
Education Dept at University of Sydney (Dr Ralph Straton) conducted a BA Honours Seminar in evaluation (students included Sue Funnell and Stephen Kemmis). This led to the first regular University courses on program evaluation.
- 1973 Karmel Report established the Schools Commission
- 1975 Adam Jamrozik, with colleagues from the Tasmanian CAE, delivers *Report on the Evaluation for the State of Tasmania*. This is the first state Government to seriously engage in evaluation for the whole of government.
- 1976 Report of the Royal Commission on Australian Government Administration (the Coombs Report)
- 1977 Evaluation and Accountability Task Force established by ACOSS, focused on developing program evaluation in the non-profit welfare agencies;
- 1978 Institute for Social Programme Evaluation established by Dr Ralph Straton at Murdoch University.
- 1979 The Baume Report (Senate Standing Committee on Social Welfare, 1979a, 1979b) *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services*.  
First issue of the *Australian Evaluation Newsletter* (produced by Preston Institute of Technology, July)
- 1982 Cain Labour Government takes office in Victoria. Institutes the Public Bodies Review Committee and Program Budgeting and gives the Department of Management and Budget the power to review existing programs in the light of the government's Economic and Social Justice strategies  
First National Evaluation Conference, Melbourne
- 1983 Review of Commonwealth Administration (The Reid Report)  
Commonwealth Government Policy Information Paper: *Reforming the Australian Public Service*
- 1984 Financial Management Improvement Program Diagnostic Study  
Victorian Evaluation Training Network formed by Jerome Winston and others;
- 1986 3rd National Evaluation Conference. Formation of the AES
- 1987 Commonwealth portfolio departments required to integrate program evaluation into corporate program management and submit evaluation plans showing systematic evaluations of all programs over a five-year cycle.
- 1989 Reform of the Australia Audit Office into the Australian National Audit Office as a result of a review by the Joint Committee of Public Accounts;
- 1990 Martin Report (Report of the House of Representatives Standing Committee on Finance and Public Administration) *Not Dollars Alone: Review of the Financial Management Improvement Programme*
- 1991 ANAO Report on Commonwealth evaluation strategy