Program Evaluation and Policy Management in Australian Central Agencies

Michael Francis Di Francesco

A thesis submitted for the degree of
Doctor of Philosophy
of
The Australian National University

December 1996

Declaration

I hereby declare that this thesis is the result of my own independent research and that all authorities and sources which have been used are duly acknowledged.

Michael F Di Francesco

Acknowledgements

This thesis has benefited greatly from the advice and assistance provided by many people. I gratefully acknowledge the financial assistance provided in the form of an Australian Postgraduate Award, as well as a postgraduate travel grant from the Association for Canadian Studies in Australia and New Zealand. The thesis would not have materialised but for the many officers of the Australian Public Service who gave so willingly and generously of their valuable time. In particular, I would like to thank Mr Keith Mackay and the staff of the Evaluation and Staffing Analysis Branch of the Commonwealth Department of Finance who graciously accommodated my intrusion and innumerable requests for information during the first half of 1995.

The Public Policy Program has been a most congenial environment in which to undertake postgraduate studies. It is a small but lively department in which I was always made to feel at home, and I would like to express my gratitude to a dedicated group of staff and students. First to my supervisory committee. Dr John Uhr, who first suggested the idea of evaluating evaluation, has been an exemplary supervisor and academic mentor. I especially thank him for his eternal optimism, good humour and superb critical capacities. Professor Richard Mulgan has cheerfully read and re-read drafts of chapters and polished what was often very rough raw material. Professor John Halligan of the University of Canberra provided pertinent advice at a crucial stage. My thanks also go to other members of the Public Policy “family” who have, at all times, provided guidance, advice or simply a kind word—Professor Francis Castles, Dr Raja Junankar, Mrs Susan Lindsay, Mrs Margot Martin, Dr Siwan Lovett, Mr Andre Moore, Ms Jennifer Curtin and Mr Mark Edwards. I extend my gratitude to my family, Anthony, Patricia and Lynda for their unerring support from afar. Lastly, but never least, I especially thank my patient wife Karon who has not known married life without *the thesis*. She buoyed me when the seas became rough and filled my sails when land, finally, was in sight.

Abstract

Of the many components of reform to Australian government administration in the 1980s, the introduction of systematic program evaluation is perhaps one of the least examined. This thesis seeks to assess the Federal Labor Government's evaluation strategy as an instrument for enhancing what are here termed the policy management capacities of central agencies. It proceeds in two steps. First, the thesis traces in detail the development of program evaluation policy in Australian federal government from the effectiveness reviews of the Coombs Report of 1976 to the current evaluation strategy, and argues that, despite competing purposes for it, evaluation was intended primarily to serve decision making in central government. This policy aim was cemented by the economic crisis of the mid 1980s and framed around budgetary issues by its steward, the Department of Finance. Second, in order to assess the impact of the evaluation strategy, the thesis develops a framework for analysing program evaluation as one instrument for strengthening the core policy management functions of central agencies. In this context, policy management is essentially a coordination task. The contribution of evaluation to two aspects of policy management—resource coordination, and policy development and coordination—is examined. The findings confirm that attempts to formalise evaluation processes have had a variable impact—central budgetary processes remain dependent on relatively informal assessment procedures, although recent attempts to enhance policy coordination through the evaluation of policy advising processes have proved potentially to be more influential. In conclusion, the thesis argues that the evaluation strategy represented a credible attempt to better inform policy making in central government, but suffered for want of clear policy design and firm execution that resulted in only a marginal impact on these processes.

Table of Contents

PART TWO: DESIGNING THE EVALUATION FUNCTION

PART THREE: EVALUATION PRACTICES

PART FOUR: CONCLUSION

Figures and Tables

FIGURES

TABLE

Abbreviations

ABC	Australian Broadcasting Corporation
ADP	Automated Data Processing
AEP	Agency Evaluation Plan
AGPS	Australian Government Publishing Service
ALP	Australian Labor Party
ANAO	Australian National Audit Office
APS	Australian Public Service
ASO	Administrative Service Officer
ATO	Australian Taxation Office
ATSI	Aboriginal and Torres Strait Islanders
BIE	Bureau of Industry Economics
CAA	Civil Aviation Authority
COAG	Council of Australian Governments
CPD	Commonwealth Parliamentary Debates
DAS	Department of Administrative Services
DASET	Department of Arts, Sport, the Environment, Tourism and Territories
DEET	Department of Employment, Education and Training
DFAT	Department of Foreign Affairs and Trade
DHHCS`	Department of Health, Housing and Community Services
DHSH	Department of Human Services and Health
DILGEA	Department of Immigration, Local Government and Ethnic Affairs
DITRD	Department of Industry, Technology and Regional Development
DoF	Department of Finance
DPIE	Department of Primary Industries and Energy
DPMC	Department of the Prime Minister and Cabinet
DSS	Department of Social Security

DTC	Department of Transport and Communications
DVA	Department of Veterans' Affairs
ECG	Evaluation Consultancy Group
ERC	Expenditure Review Committee (of Cabinet)
ESAB	Evaluation and Staffing Analysis Branch
ESSB	Evaluation and Statistical Services Branch
ESU	Efficiency Scrutiny Unit
FESP	Forward Estimates Strategy Paper
FMIP	Financial Management Improvement Program
GDP	Gross Domestic Product
GED	General Expenditure Division
GFEP	General Financial and Economic Policy Division
HRSCE	House of Representatives Standing Committee on Expenditure
HRSCFPA	House of Representatives Standing Committee on Finance and Public Administration
HRSCTCI	House of Representatives Standing Committee on Transport, Communications and Infrastructure
IDC	Inter-departmental Committee
JCPA	Joint Committee of Public Accounts
JMR	Joint Management Reviews
MAB	Management Advisory Board
MATS	Modernised Apprenticeship and Traineeship System
MBS	Medical Benefits Scheme
MIAC	Management Improvement Advisory Committee
NPP	New Policy Proposal
OCG	Office of the Comptroller-General of Canada
OPAAM	Office of Policy Analysis and Administrative Management
OSC	Commonwealth-State Officials Steering Committee
PAR	Program Analysis and Review
PEAC	Program Evaluation Advisory Committee
PEMS	Policy and Expenditure Management System
PEP	Portfolio Evaluation Plan
PES	Public Expenditure Survey
PEU	Program Evaluation Unit
PMB	Program Management and Budgeting
PMPR	Program Management Performance Reviews
PMR	Policy Management Review
PPS	Program Performance Statement
PSB	Public Service Board

PSC	Public Service Commission
RCAGA	Royal Commission on Australian Government Administration (The Coombs Commission)
RIPAA	Royal Institute of Public Administration Australia
RMI	Resource Management Improvement Branch
RNSWGA	Review of New South Wales Government Administration
SAC	Structural Adjustment Committee (of Cabinet)
SO	Savings Option
SOG	Senior Officer Grade
SSCFPA	Senate Standing Committee on Finance and Public Administration
SSCSW	Senate Standing Committee on Social Welfare

Chapter One

INTRODUCTION: PROGRAM EVALUATION AND AUSTRALIAN PUBLIC POLICY

> Here is an infallible rule: a prince who is not himself wise cannot be well advised . . . when seeking advice of more than one person a prince who is not himself wise will never get unanimity in his councils or be able to reconcile their views. Each councillor will consult his own interests; and the prince will not know how to correct or understand them. Things cannot be otherwise, since men will always do badly by you unless they are forced to be virtuous. So the conclusion is that good advice, whomever it comes from, depends on the shrewdness of the prince who seeks it, and not from the shrewdness of the prince on good advice (Machiavelli 1995 [1961], ch xxiii, pp75-76).

On the virtues of undiluted political rationality in the public policy process, students of politics can find no better exponent than Niccolo Machiavelli. Writing in a time of immense political intrigue and tumult within the declining Italian city states in the early sixteenth century, Machiavelli the civil servant was centrally concerned with the practicalities of strong (if not ruthless) government that endured. Within an increasingly fractious polity, good government—and, by extension, good public policy—would always require two symbiotic elements: effective political leadership and policy advice (or knowledge) on which to act. However, as the epigraph above reminds us, effective political leadership required an uncommon mix of prudence and resolve, for advice can often serve interests other than those of the prince and it was the task of the ruler to judge both the veracity and appropriateness of the options placed before him. For Machiavelli, this relationship between leadership and advice ran strongly in one direction—the value of any policy advice could, ultimately, only be judged with reference to the political acumen of the prince.

Machiavellian pronouncements on statecraft, although “an impassioned tract for the times” (Machiavelli 1995, p(x)), resonate with one of the perennial questions facing modern governments—how to secure the best policy advice available. The story of government and its advisers in Australia over the last two decades is certainly no exception. In a period of prolonged economic and political disorientation throughout the late 1970s and 1980s, governments at both federal and state levels have sought to improve the quality of the policy advice they receive through two key developments (Halligan & Power 1992). The first hinged on increasing political control of the bureaucracy by *breaking* its monopoly on policy advising. The chief instrument was the appointment of well resourced ministerial advisory staffs. The second was framed around a fundamental restructuring of the principal existing source of advice on public policy—the public service. There remained, of course, a presumption that the government (in the form of the *political* executive) was possessed of sufficient political acumen. The most important task was to secure “better” policy advice, particularly at that level of the public service machinery that serviced cabinet decision making. The subject of this thesis comprises a central aspect of this second task. It is concerned with assessing the processes and outcomes of administrative restructuring as they relate to one important component of the policy advice system—program evaluation and the policy advising capacities of central agencies.

Aim And Scope

A prominent text on public policy in Australia has noted how, in contrast to international practices, public policy studies in the antipodes have developed within a markedly more pragmatic tradition of case and institutional studies, with the consequence that “the literature is structured around topics rather than around competing paradigms—there are theories within the field, rather than theories of public policy” (Davis, Wanna, Warhurst & Weller 1993, pp11-15). This more utilitarian approach has seemingly weathered the rise of reputedly more theoretical analytical tools, such as the public choice school and the so called “new institutionalism”, precisely because it already stressed the importance of political institutions (such as political parties, parliament, the public service, and diverse corporate interests) in shaping the outcomes of public policy choices. According to the same authors:

> Policy and politics are not easily separated, since each informs the other. But neither should policy simply be reduced to consideration of the politics of the moment. History and process matter, as do the values, interests and resources competing in any policy arena. All influence outcomes, and require a place in analysis (Davis, Wanna, Warhurst & Weller 1993, p16).

This represents something of a 'holistic' approach to analysing public policy, combining history, interests and resources of identified political institutions, as well as an examination of the policy process, in tracing the determinants of policy outcomes. This general approach will be adopted here to frame an examination of the development of a program evaluation policy within Australian federal government.

For Australian public policy the 1980s could easily be characterised as the decade of administrative reform in government. At the federal level successive Labor Governments, first under the stewardship of Bob Hawke and then later Paul Keating, initiated and implemented a sweeping agenda of public sector restructuring framed specifically around reforms to systems of financial management and accountability. A prominent component of these changes was the introduction in 1988 of a new public service-wide policy of program evaluation designed to assist both political and bureaucratic decision makers by providing information on the efficiency, effectiveness and appropriateness of government programs. However, like so many of the individual components of the Labor Government's public service reform program, the "evaluation strategy" has not yet been subject to a satisfactory level of external scrutiny and academic analysis (cf Halligan, Beckett & Earnshaw 1992, p7; Pollitt 1995, pp135, 138-140). The little that does exist can be divided into two categories. The first benignly takes the rationale for program evaluation—as a tool for assisting the implementation and management of government programs—at face value, applauding its focus on the effectiveness and relevance of government policies but cautioning policy designers to take account of the role that politics inevitably plays in any evaluation process (Uhr 1990b, pp89-90; Uhr & Mackay 1992, pp433-434). The second group is somewhat more critical, claiming that program evaluation appears to be nothing more than a recycled tool of central agency control that runs not along the new effectiveness grain, but against it by remaining locked into issues of economy (Ryan 1992, pp71-72; Caulley 1993, pp3-7; Colebatch 1995, pp150-151, 156-161).

Holding up the eye of doubt, this thesis takes its cue from the second school, and seeks to assess the evaluation strategy as an instrument for both preserving and enhancing what are here referred to as the policy management capacities of central agencies. At its base, we can take program evaluation to be the production of information for decision making by specified institutional policy actors. This prompts two key questions. For whom is the information produced? And why? The thesis approaches these issues in two steps. First, it argues that these questions can be answered through an analysis of the demand by competing institutional interests in Australian government for program evaluation. These interests would ultimately determine the final shape of the current policy. While the evaluation strategy has been prosecuted by the Department of Finance in the name of the Labor Government, it is perhaps more accurate to depict stewardship as residing with that department. Certainly there was strong ministerial support for the policy, but this was more often than not a reaction to policy initiatives on the part of Finance which had, since July 1987, eagerly assumed responsibility for management reform. The onset of fiscal stress during the mid 1980s forced the Labor Government to adopt “decremental budgeting” strategies and within this climate of public sector contraction the Department of Finance seized on program evaluation as a multipurpose instrument—not only did it serve the most pressing need for identifying expenditure cuts, it also helped secure the place of the budget agency within an administration undergoing comprehensive restructuring.

However, as convenient as this immediate policy response may have appeared, the evaluation strategy was not solely a product of its time; it was also a captive of its own policy history. While the interplay of a number of institutional interests saw competing purposes for program evaluation—as a tool for program management, parliamentary accountability of the executive and central decision making, which were all incongruously reflected in the evaluation strategy—evaluation policy for *central coordination*, a view first countenanced by the Coombs Report of 1976, remained the strongest current. The stream of policy development from the mid 1970s through to the evaluation strategy has ensured that program evaluation would always be the servant of the coordination needs of central agencies.

In light of this, the thesis seeks to do two things. The first is to provide a detailed historical commentary on the development of a program evaluation policy within Australian federal government. This will show that despite the competing purposes for program evaluation, the prevailing intention was for it to serve the information needs of central agency (and ultimately cabinet) decision making. During the heady reform days of the mid to late 1980s, program evaluation was driven predominantly by the information needs of the central budget agency, the Department of Finance. Second, following on from this analysis, the thesis seeks to assess the outcomes of the Labor Government's evaluation strategy against these primary, although somewhat latent, objectives in terms of its contribution to the policy advising capacities of the central agencies servicing cabinet (the policy management function). In particular the impact of the evaluation strategy on two aspects of policy management—resource coordination and policy development and coordination—are assessed. Even though the strategy appears to have been retained by the new Howard Coalition Government, the election of that government in March 1996 presents a convenient end marker for the study. To this extent the thesis is a post-mortem of a narrowly defined policy initiative.

DEFINING PROGRAM EVALUATION

As already suggested, this study examines the introduction of a specific analytical technique—program evaluation—into the decision making processes of central government. In sum it is centred on the *politics* of this development which can, for some public policy analysts, be reduced to efforts to move evaluation information "from those organisations that have it to those that need it" (Peters & Barker 1993, p5). There is, of course, an almost overwhelming literature, sourced predominantly from the United States, that deals with program evaluation and the application of social science to government and public policy. As one way of further demarcating the scope of this study, it would be prudent, at the outset, to define what is meant by the term "program evaluation" and, at the same time, state what is *not* canvassed by this examination.

Earlier, it was noted that program evaluation could be seen quite simply as the production of information for decision making by specified institutional policy actors. What defines program evaluation is the *form* this information takes. There are, of course, a multitude of ways of conceptualising the types of information which are produced specifically for public policy decision making, the most common being framed around the generic term "research" and its utilisation (Dunn & Holzner 1988, pp5-9, 12-13, 15-17; Dery 1990, pp6-25; Lester & Wilds 1990, pp313-316). In this thesis, the preferred approach will be to categorise information for government decision making into three broad types—data, research and analysis (Lindquist 1988, pp88-89).

The most basic type of information designed to influence decision makers is data. Data refers to the "mechanical processes" of generating and collecting what can be referred to as "factual" information, such as broad economic and social statistics as well as more specific indicators of the performance of government programs, which require interpretation and may inform both research and analysis (cf Peters & Barkers 1993, pp9-11). The second type of information is research. This refers to the careful consideration or study of a set of policy conditions in order to understand the underlying relationships between different variables. In the North American literature on evaluation, program evaluation is often equated with what is termed "evaluation research", or the "process of applying scientific procedures to accumulate reliable and valid evidence on the manner and extent to which specified activities produce particular effects or outcomes" (Mayne & Mayne 1983, p273). This definition reflects the origins of evaluation in the United States during the mid 1960s as a tool specifically targeted at improving the administration of social programs through the application of the aims and methods of social science research.

The final type of decision information is analysis. In advancing his typology Lindquist argues that there are significant differences between research and analysis:

> Analyses can be reasonably lengthy, contain careful argument, and use much data; but it is a limited enterprise, constrained by time, and intended to produce specific recommendations. Analysis takes its sources of data and the perspectives of disciplines as given (Lindquist 1988, p89).

Analysis then refers to information that may be produced by research-type methods, but is specifically modified to fit the resource and time constraints of the decision making process in government—it is "more policy relevant and decision-oriented" (Mayne & Mayne 1983, p275). In this thesis, program evaluation, as defined by the Labor Government's evaluation strategy, will be taken to be a form of policy analysis. As the application of social science techniques, program evaluation may use either existing data sources or social science research methods, or both, but it is, unlike these two earlier instruments, specifically designed to provide options for *constrained decision makers.*

A consideration of program evaluation as analysis will inevitably buy into the continuing debate on the *utilisation* of social science derived information in government decision making. In this there are essentially two positions (Davidson & White 1988, pp1-5; Dunn & Holzner 1988, pp15-17). The first questions the value of rational decision making by claiming that there is no direct or observable link between information and policy making (or, at the very best, if a link exists it is contingency based). The second contends that utilisation is indirect and certainly less measurable since information "enlightens" policy making over time by influencing climates of opinion. While any attempt to assess the outcomes of a program evaluation policy must give some consideration to these questions of utilisation, the focus of the present study is less on how program evaluation is used in final decision making on discrete policy issues and more on how it contributes to the formulation of central agency policy advice; that is, evaluation's influence on the procedural aspects of policy management. In this the study presumes the potential for both types of information utilisation. This has two further consequences that delimit the focus of the present examination. First, the thesis does not purport to analyse the epistemology of the role of social science in public policy making (Albaek 1995), nor attempt to contribute to the continuing debate on the merits of competing "paradigms" for structuring social inquiry (Guba 1990, pp18-27). Second, it does not seek to explore the nature of program evaluation as a professional pursuit, nor attendant questions about competing evaluation methodologies, although both are likely to have some influence on the way an evaluation policy is designed (Chelimsky 1989, pp12-21).

METHOD OF INQUIRY

At a general level the reforms to public sector management in the commonwealth government over the past thirteen years have been exceptionally well documented, both in terms of publications produced by government as well as commentary contained in secondary sources (eg. MAB-MIAC 1993a; Campbell & Halligan 1992; Halligan & Power 1992; Zifcak 1994).

However, as noted above, there remains a gap in terms of analysis of specific reform initiatives. This has particular application to program evaluation, which only approached full implementation towards the end of the third Labor Government. Since the research undertaken for this thesis was directed at the way in which program evaluation had influenced the *current* policy management functions of central agencies, there was always going to be a question mark over access to sufficient sources of information. A number of comparative studies of administrative behaviour in central agencies have noted how research constraints can tighten when access to both written and oral evidence is restricted by busy government departments (although the Canberra bureaucracy is generally perceived to be both less insular and more open than comparable administrations overseas) (Campbell 1983, pp357-359; Campbell & Halligan 1992, pp238-240; Zifcak 1994, pp197-199).

Within these constraints the research for the thesis relied on three principal sources of information. The first comprised participant interviews within the relevant government departments. The second consisted of information gathered from examination of government documentation. This included both internal departmental files, archival material and external, publicly available publications. The third source was secondary material in the form of commentary and analysis contained in books and professional journals.

Interviews

There are a number of approaches that researchers can employ when undertaking interview research in administrative settings. For instance, the researcher could adopt an anthropological approach which relies on a ‘participant-observer’ role and working alongside public officials for an extended period (Kaufman 1981, pp4-13). Another approach relies on more circumspect entry to the bureaucracy through what has been termed the “process of referral” (Zifcak 1994, pp197-199). This involves building up interviews with officials *solely* by way of referral from someone connected with the bureaucracy who was already known to the researcher. This second method was adopted for the thesis. It was initiated by approaching the secretary to the Commonwealth Department of Finance, who was known to the author’s academic supervisor, and from whom came official approval for the author to undertake research in the Evaluation and Staffing Analysis Branch

(ESAB) of the Department of Finance during the period March-July 1995.[1] Using ESAB as a base, interviews were secured in other branches within Finance as well as other departments. Because of similarities in the design of evaluation policies, this Australian material was supplemented by a small number of interviews conducted with senior officers within the Canadian federal government during September-October of 1995.

The interview program was structured around the research focus on central agencies, and, in general, excluded the perspective of line portfolios. This approach could be charged with prejudicing the study's findings, but it was adopted nonetheless so as to ensure a manageable research topic. The interview program had two primary goals—first, to determine how the evaluation strategy was planned and implemented, and second, to gain some understanding of the respondents' attitudes to the impact of this development. As indicated above, interviews were arranged within the Department of Finance on a referral basis as the author familiarised himself with officers located in different branches of the department. The vast majority of interviews within Finance were conducted with both senior (secretary, deputy secretary and assistant secretary level) and middle (ASO6 to SOGA level) management officers located within both ESAB and three of the "supply divisions". Whilst located in Finance, the author also secured interviews with senior officers in other central agencies (Departments of the Treasury and Prime Minister and Cabinet) and, where pertinent, line departments (Department of Human Services and Health).

A total of thirty interviews, each of which averaged approximately one hour, were conducted, twenty-four in Canberra and the remainder in Ottawa. The interview guide followed a non-standardised or unstructured format in which a list of topics was used to sequence predominantly open-ended questions (Fielding 1993, pp136-139; Patton 1990, pp280-284). To this extent the interviews followed a relatively informal discussion style. This open format provided the interviewer with considerable scope in phrasing questions to suit the interview conditions (for instance, the office and seniority of the respondent) and pursuing related issues of importance as they arose. Interviews were conducted on the basis that interviewee's responses were attributable unless otherwise requested. In only a very small minority of cases was permission withheld. However, because of the sensitivity of the material,

[1] A list of all the abbreviations used throughout the thesis appears at pp(vii-ix).

in terms of both potential prejudice to individual officers and its relationship with current policy, it was decided subsequently to mask direct attribution of interview material. Interviews have been coded and direct attribution—in the text, at least—avoided. This also follows the standard practice in policy studies utilising interview material (Campbell 1983, pp357-361; Campbell & Halligan 1992, pp240-241; Savoie 1990, pp20-21; 1994, pp14-15; Weller & Grattan 1981, pp17-21; Zifcak 1994, p201). The list of interviews appears at Appendix I.

All interviews were recorded manually, with the transcripts—which were prepared soon after interviews—being as thorough and comprehensive as such note taking permits. Even though the tape-recording of interviews is much the preferred method of transcribing non-standardised one-on-one interviews (Fielding 1993, pp145-147; Patton 1990, pp347-353), the taping of interviews conducted with *current* public officials can sometimes have the effect of reducing the openness of interviewees. (Further, to the extent that there may have been sensitivities relating to the subject matter of questioning, interviews were aided by a non-standardised format.) The position taken here was that manually recorded interviews, though they run the risk of reducing the accuracy of interview transcripts, possessed the twin advantage of maximising the potential for a frank discussion and minimising the temptation on the part of the researcher later to rely excessively on direct quotations as stand-alone evidence.

As in any social research utilising interviews, there can be problems in interpreting data. It is best to concede these limitations from the outset. First, from above, the reliability of data sourced from untaped non-standardised interviews can be open to some doubt—non-standardised interview techniques can attract accusations of interviewer bias in leading interviewees (Fielding 1993, pp140-141, 147-150). Second, it is not unknown for interview subjects to adapt their responses to the very fact that they are being interviewed (for instance, according to what the interviewee thinks the interviewer *expects* to hear) (Kaufman 1981, p13). Third, the potential for falling victim to the “fallacy of composition” must also be acknowledged. Since interviews could not be conducted with each and every officer within, for example, the supply divisions of the Department of Finance, care must be taken in interpreting evidence such that what is true for some officers and branches is not *assumed* to be true for all. With the usual proviso relating to these limitations, it is submitted that the data derived from the interviews is nonetheless instructive in helping to explain the impact of program evaluation on the work of central agencies.

Internal Departmental Files

In the task of piecing together the jigsaw of bureaucratic policy development, interview data, as the preceding discussion indicated, represents only part of the picture. One must begin with the written record. Generous as the Department of Finance was in granting entry, there were restrictions on the internal files which could be accessed (most obviously in relation to those classified ‘cabinet-in-confidence’). However, in most instances approval was given to see a large proportion of the material relating to the development, implementation and progress of the Labor Government’s evaluation strategy. Although the photocopying of material was not permitted, there were no restrictions placed on the taking of hand written notes. Given the imposing volume of file material, this condition, combined with the tight schedule for the interview program (the timing of which coincided with the lead-up to the 1995 Federal Budget), ensured that some selectivity was required in examining departmental files. This may have prejudiced subsequent interpretation.

External Documentation

It need hardly be said that there are dangers in relying solely on source material emanating from a single agency, in this case the Department of Finance. The principal concern is whether or not the researcher is seeing and hearing what the agency wants him to observe rather than that which he *could* be exposed to. This hazard is perhaps even more pronounced for an outsider entering an organisation where he or she may be more prone to accepting the “departmental line”. In order to guard against this kind of capture, every effort has been made to utilise publicly available review documentation as an aid in the task of examining Department of Finance material. Ultimately, along with interviews conducted outside of Finance, this provides the only source of evidence external to the bureaucracy that can be employed to assess the outcomes of the evaluation strategy in terms of the policy management functions of central agencies. In particular, reports emanating from the Auditor-General and relevant committees of the Commonwealth Parliament were invaluable. Performance information contained in departmental annual reports, program performance statements and general public service reviews also proved to be useful.

STRUCTURE

The thesis consists of an introduction, and seven chapters divided into four substantive parts. Part One begins by constructing a framework for analysing the introduction of a formal program evaluation policy within Australian federal government. Chapter Two recognises that the use of policy analytical techniques in government is an international phenomenon and seeks to locate Australian developments within this wider context. It concentrates on the processes of and the reasons for the institutionalisation of program evaluation within the governments of industrialised countries, and contends that the increasing use of evaluation has been driven by a combination of economic stress and the perceived need to restructure the machinery of central government. While the Australian case generally conforms to international patterns of development, it is also distinctive because public sector restructuring was tempered by the specific policy commitments of the ruling Australian Labor Party.

Having made a case for narrowing down the focus to central agencies, Chapter Three develops a framework for analysing the impact of program evaluation on the work of these agencies. The framework consists of two propositions. First, that the core functions of central agencies can most profitably be framed around the management of the coordination problem in central government, here referred to as policy management. Second, that these policy management capacities have been enhanced in response to the transformation of the coordination problem from an initial concern with the process of government to a sharpened focus on substantive policy concerns, and that program evaluation has been a crucial instrument for this transition. This study seeks to test the validity of the second proposition by asking two key questions. Has evaluation recast the way that the policy management function is approached? Does evaluation lead to the over-extension of central agency policy and resource coordination capabilities?

Part Two begins this task by providing a detailed historical account of the development of program evaluation within Australian federal government, and confirms its role as principally an information tool for central government.

Using the accountable management framework initiated by the 1976 Coombs Report as a point of reference, Chapter Four traces up to 1986 the stream of policy development that influenced the current evaluation strategy. Representing the new review criterion of effectiveness, program evaluation proved to be a fiercely competitive policy arena in which different institutional interests sought to impose a functional role. Three competing purposes emerged—evaluation for central policy management, evaluation for informing policy implementation, and evaluation for external accountability. Although evaluation for central policy management would prevail, not least as a defensible role defining the newly created Department of Finance, there remained no commonly accepted functional role for program evaluation in government.

The historical analysis concludes in Chapter Five which follows the design and implementation of the evaluation strategy after 1987. Evaluation as an instrument for central policy management is cemented in a period of acute fiscal stress as stewardship is finally handed over to the central budget agency, the Department of Finance. The evaluation strategy emerged as a key component of public service financial management reform and principally as a tool for imposing fiscal discipline, although it was promoted as an uneasy amalgam of the three competing objectives that surfaced in the early design period. However, because it purported to serve a range of purposes, the evaluation strategy faced implementation problems that threatened the accomplishment of all of them, policy management included.

Part Three addresses the performance, both actual and promised, of the evaluation strategy with respect to the policy management responsibilities of central agencies. It is structured around two dimensions of the policy management function, resource coordination, and policy development and coordination. Chapter Six examines the impact of program evaluation on the central budgetary process. Specifically, it seeks to assess the degree of change in the approach of budget officials (or supply officers) to their resource coordination responsibilities and the extent to which evaluation has influenced the core task of the budget agency—the provision of policy advice relating to the expenditure priorities of government. By attempting to formalise aspects of the budgetary process that are intrinsically informal in nature, the evaluation strategy failed to effect enduring change in policy management functions. Contrary to the conventional (Department of Finance) wisdom, the main finding is that evaluation’s impact on resource coordination has been marginal.

The focus of Chapter Seven is the contribution of program evaluation to policy development and coordination. The evaluation of policy advising programs, which was intended to inform both public accountability and policy development objectives, was the last and most controversial phase of the evaluation strategy, heralding the introduction of a distinct form of evaluation, the policy management review (PMR). The evaluation of policy advising programs was a departure from the strict effectiveness focus of the evaluation strategy in that it signalled a return to assessing (interdepartmental committee) policy *processes*. There are two main arguments. First, that the PMR format was the product of bureaucratic expedience—evaluating the effectiveness of advisory processes was shaped by a combination of methodological constraints and strong bureaucratic opposition. Second, that the PMR format has emerged as a tool for increasing the influence that central agencies have in the processes that generate policy advice. The chapter analyses the PMRs completed to date and concludes that the evaluation of policy advice represents the most express application of evaluation to policy management purposes.

The study concludes with Chapter Eight. This draws together the findings of the two case studies and examines the implications of their results for the policy management functions of central agencies. It presents an overall assessment of the evaluation strategy as policy design and points to the uncertainty of political choice as the crucial factor conditioning the introduction of any policy analytical technique. While the evaluation strategy can be adjudged a credible attempt to systematise policy evaluation—hampered as it was by flaws in both policy design and implementation—in light of the constraints imposed by political choice, the evaluation strategy, like so many policy analytical reforms before it, can also be seen as another case of government seemingly ploughing the policy sands.

Chapter Two

Policy Climate, Evaluation And The Reshaping Of Central Government

This chapter begins by making a case for the examination of central government departments as the crucial agents driving program evaluation policy in Australian government. It argues that the demand for program evaluation by Australian government in the late 1980s can best be explained as one response to a distinct set of economic and political conditions. This response was framed around a fundamental restructuring of both government and economy. To this end, the chapter also sets out to chart the policy climate prevailing in Australia at this time, as well as sketch the distinctive response of Australian policy makers.

The chapter proceeds in two steps. The first section is a survey of the system-level explanations for the institutionalisation of program evaluation within government over the last twenty years. This comparative literature argues that the most important determinant of this process is change in the *nature* of demand by governments for program evaluation. This demand is a function of the prevailing policy climate and it is only through the creation of policy corridors—a confluence of favourable policy conditions—that program evaluation is adopted by governments. Persuasive as this general explanation is, the principal argument of this chapter is that it has only partial application to the Australian case where a unique constellation of political, economic and bureaucratic factors generated a *distinctive* program evaluation policy. To this end, the second section examines the nature of the policy corridor that opened in Australia. It argues that the corridor was framed around the issues of economic and governmental restructuring. While the political imperatives of the ruling Australian Labor Party (ALP) conditioned responses to these issues, this section reinforces the common view that central economic departments played a key role in reshaping the machinery of central government.

THE EVALUATION FUNCTION: COMPETING EXPLANATIONS

The last thirty years have witnessed efforts by most western industrialised countries to impose greater rationality on their policy making capabilities, most prominently through the techniques offered by policy analysis. Program evaluation is one technique that has been adopted with remarkable consistency during this period, and with renewed verve in the last decade. The aim of this first section is to place recent Australian developments in an explanatory context, by surveying that comparative literature which seeks to identify trends in the application of program evaluation to government. This literature is marked by widespread agreement that the most significant factor driving the development and use of program evaluation has been changes in the demand by government for this type of analysis. More specifically, these patterns of demand have been shaped (or created) predominantly by central agencies in response to fiscal stress. While accepting the import of these general patterns, this section contends that the observed trends have only partial application to the distinctive Australian case.

In the last decade a growing body of cross-national research has emerged that seeks to explain the development of the program evaluation function within parliamentary and presidential (separation-of-powers) systems of government in North America and Western Europe. This literature concentrates on the issue of *institutionalisation*—that is, why governments adopt program evaluation and how they have integrated it into policy making processes. This type of comparative analysis is most commonly conducted along the dimension of legislature-executive relations. Different institutional rules of policy making ensure that power differentials vary across different systems of government—in parliamentary systems legislative power is centralised with the executive through cabinet and accountability is concentrated in the legislature; in presidential systems the converse holds true (Weaver & Rockman 1993, pp11-16; cf Levine 1981, pp33-35). As a consequence, the executive tends to be less constrained in (majoritarian) parliamentary systems than in presidential style systems. The institutionalisation of program evaluation reflects this. In parliamentary systems program evaluation is overwhelmingly an executive instrument, while in presidential systems it is pursued by both executives *and* legislatures, as legislatures have sought to increase their influence within more diffuse power structures (Derlien 1990, pp155-157; cf Mosher 1979, pp174-181). Although broad trends will be traced here, our concern primarily is with parliamentary systems of government.

Derlien (1990, pp149-161) has suggested that the historical process leading to the institutionalisation of program evaluation over the last three decades can be seen as a two stage model consisting of successive “waves” of what he terms the “evaluation movement”. The first wave occurred during the international economic buoyancy of the mid-to-late 1960s. It has been closely associated with two key developments. The first of these were the ambitious social reform policies introduced by “social-liberal” governments that were grounded in the conviction that the application of social science to public policy could help ameliorate social problems, particularly poverty and poor education, health and housing conditions. Framed around “social experimentation” program evaluation was intended to test the impact of social policies on targeted populations and then assist in program improvement (Shadish, Cook & Leviton 1991, pp20-25). The second concurrent development was an attempt by governments to introduce greater rationality in decision making through policy planning and program budgeting. This dovetailed with the thrust of reform programs since, in a favourable economic climate, program administrators sought the results of evaluation to justify government intervention (Hellstern 1991, pp272-273). The first wave, which had its origins in the United States of America, is characterised by a pronounced leaning towards evaluation and program administration.

The second wave of evaluation can be traced from the late 1970s and early 1980s as a response to acute economic stress. According to Derlien, a commitment to evaluation was renewed by “predominantly, although not exclusively, conservative governments” convinced of the pressing need to restrain government spending (1990, p161). Drastically altered economic and political climates placed new limits on government intervention and attached a high premium to the improvement of government efficiency. The new policy conditions had two effects. First, fiscal stress redirected the focus of program evaluation from the established concerns with policy impact to an occupation with the bare economy or efficiency of programs (Gray & Jenkins 1989, p24; Henkel 1991, pp17-20). Second, across all systems of government, fiscal stringency and efficiency oriented evaluation was being driven by a number of institutional actors, including legislatures, audit offices, peak cabinet structures and the central finance agencies. The dire nature of the political priorities imposed by economic conditions ensured that cabinet and central finance agencies placed great store in the budgetary implications of program evaluation (Derlien 1990, pp163-164). Most commonly, the United Kingdom, Canada and the Nordic countries are associated with this second wave.

This two-stage comparative survey is useful in briefly describing the main markers of the development of program evaluation in government. However, as Derlien (1990, p147) correctly points out, national developments are a function of “a number of factors and constellations” and it would be more instructive to look at the institutionalisation issue at a much more specific level. The preferred approach in the literature—and the one I will follow here—is to isolate those factors which influence the demand for and supply of program evaluation techniques.

Perhaps the most obvious factor influencing the demand for program evaluation in government are policy shifts caused by changes in “political regimes” (Wagner & Wollmann 1986, pp205-207; cf Gray & Jenkins 1989, p21). Here, a political regime refers to a “change in the political party dominance and ideological complexion of the central or federal government”. The regime shift argument holds that the activities of government are primarily shaped by the policy objectives of the incumbent political party. In this way, the two waves of evaluation can be explained by shifts in political party dominance from conservative to social democratic (first wave) and from social democratic back to conservative (second wave). Prime examples include both the United Kingdom and the United States in the early to mid 1980s where a switch to strong conservative governments, deeply suspicious of the connections that social science research had with the policies of the expansionary welfare state, ensured that evaluation was “ideologically stigmatised” and redirected solely at questions of cost-efficiency and management improvement (Wagner & Wollmann 1986, pp213-214). However, according to the same authors, regime shifts of this type present only part of the demand story, since successful policy change requires favourable structural conditions that create a “specific corridor of action within which reform minded policy makers act” (Wagner & Wollmann 1986, p210). Not surprisingly, a corridor is often formed by the prevailing economic conditions and the characteristics of the politico-administrative system (as well as fluctuations in the policy issue attention cycle, cf Tarschys 1983, pp377-383).

Analysing the demand for evaluation as a function of propitious economic and political conditions is a useful exercise (despite the anomalous fact that regime shifts in the Australian case, as we will see, operated in reverse). Others have taken up this theme in categorising Derlien's historical factors as either demand-side or supply-side determinants (Bemelmans-Videc 1992, pp7-12). At the general level, as both Derlien and Wagner & Wollmann contend, demand is determined by the prevailing economic and political conditions that are likely to alter the orientation of administrative structures. During the second wave, climate changes were accompanied by the transposition of a private sector derived management framework onto bureaucracies as part of the new efficiency focus. Supply-side determinants, on the other hand, comprise the availability of and, within government, the receptivity to the various types of social science skills. As Bemelman-Videc frames the term, supply is, to a considerable degree, a function of demand since often the receptivity to social science analytical skills is heavily dependent on the administrative culture prevalent in bureaucracies; for example, the legal-administrative culture of continental European countries (Bemelmans-Videc, Eriksen & Goldenberg 1994, pp150-154, 178-180) or the closed generalist tradition so dominant in the United Kingdom (Bulmer 1988, pp28-30) have both proved resistant to the influx of social science trained administrators. At the same time, supply is also determined by the level of professionalisation attained by those who claim to be practitioners of program evaluation (Bemelmans-Videc 1992, pp13-15; Hellstern 1991, pp277-278).

The contention here is that the focus should be on demand-side developments. This does, of course, ignore the dynamics of interaction between supply and demand, since the relative contribution each makes to institutionalisation remains open to debate. Certainly, there is some evidence suggesting that certain "policy booms" in government have been generated by the confluence of corporate and professional interests with emerging political priorities (for instance, privatisation policy; see Dunleavy 1986, pp13-15, 30-32). In these cases, private sector corporate and professional interests "market" the benefits of new policy techniques for government and push for their adoption by public agencies. Commonly the marketing campaigns have secured prior government approval. The ready adoption by governments of private sector management techniques (including variants of program evaluation) during the second wave can, to some extent, be explained as part of the 'solution packages' pushed by burgeoning management consultancy interests. However, as suggested above, a supply-side argument cannot, by itself, provide a plausible explanation as to why program evaluation activity has intensified.

In contrast, the value of a demand analysis of program evaluation in government is that it allows us to narrow down the effects of policy environmental changes to a change in the *nature* of demand. The explanations for institutionalisation reviewed thus far have suggested that new policy imperatives precipitated altered bureaucratic structures and ultimately the primary task facing evaluation. In effect, the demand for evaluation shifted constituencies, moving from the domain of program administration and policy implementation to the budgetary requirements of spending conscious cabinets and central finance agencies. The demand for program evaluation was being driven by a reshaping of administrative and policy priorities:

> Typically, the bearers of this second evaluation movement are not the program administrators in the government departments, but rather the finance ministers and the auditing offices as the traditional wardens of the budget. Consequently, the perspective on and the function of evaluation slightly shifted: instead of effectuating programs, the emphasis was rather on curbing ineffective programs in order to cut back the national budget (Derlien 1990, p154).

'Perspective', however, may describe this functional redesignation a little too benignly. The role change has been consciously engineered by central review agencies which, in many countries, have been charged with the responsibility of acting as "change agents" within government (Derlien 1990, pp164-165; Bemelmans-Videc 1992, p11; cf Levine 1981, pp29-30, 59-60). Although some commentators have observed considerable variation in the extent to which responsibility for program evaluation is centralised within governments (Rist 1990, pp7-9), it is fair comment to conclude that the transformation of demand has induced a new orthodoxy of program evaluation as a tool for expenditure control.

Changes in the complexion of demand have also been the theme of the most recent treatments of evaluation (Gray, Jenkins & Segsworth 1993; Gray &

Jenkins 1992, 1993). Approaching the question of demand from a political systems analysis these studies arrive at much the same conclusions. They see program evaluation as one of three “processes”—budgeting, auditing and evaluation—that serve a number of functions necessary for the maintenance and performance of any political system. This functional schema suggests that the emergence of program evaluation is inextricably linked with reforms to central budgetary processes in response to external economic pressures (Gray & Jenkins 1993, pp190, 206). The argument is that the new environmental norm of fiscal stress has accelerated efforts by political systems to attain the objective of a “mature” budgetary system whereby the three processes, and the various functions they perform, are fully integrated. The level of integration is determined by the satisfaction of three “preconditions” or what can be here termed measures of demand—the demand for new technical skills, the demand or acceptance of evaluation within existing decision making procedures, and the demand for re-examination of established policy norms by the relevant political authority (Gray & Jenkins 1993, pp200-204; 1992, 65-67).

Functional integration as outlined in the preceding paragraph is a normative ideal. After surveying the place of evaluation in the governments of seven industrialised countries (excluding Australia), Gray & Jenkins conclude that “no country is mature in this sense [of full integration] or even moving toward such maturity” (1993, p199). For them, the failure to integrate could be attributed to one intractable problem—despite moves to link it with central budgetary processes, program evaluation continued to serve a functional role within government that failed to “establish clearly whom (or what) it is supposed to serve” (Gray & Jenkins 1993, p190). The consequence is that program evaluation has generally been misapplied as a tool serving diverse management purposes—resource allocation, program implementation and both internal and external accountability—rather than being directed solely at central questions about policy and budgetary priorities.

As Gray & Jenkins note, the absence of budgetary integration is also supported by Schick’s survey of budgetary adaptation in industrialised countries (this time including Australia) (Schick 1986a, 1988, 1990). Schick contends that comprehensive budgetary reform has been replaced by “adaptation” at both macro and micro-levels in response to deteriorating economic circumstances. At the macro-level, where total expenditure is considered, the budget process has been altered so as to “dampen the pressure for expenditure and to strengthen the capacity for restraint" (Schick 1986a, p214). Here, the shift in orientation from planning tools to expenditure control has been achieved via increasing use of tight fiscal targets and multi-year budgets or forward estimates. At the micro-level, where central finance agencies apply limits to specific program allocations, adaptations have, according to Schick, emerged as "small adjustments" (Schick 1988, p532). And here program evaluation is a case in point. For while governments had gradually recognised that program evaluation was most effective when it "comes through the recurring procedures and pressures of budgeting, not through special activities", Schick concludes that cutback management did *not* provoke governments into establishing formal systems of program evaluation (Schick 1988, pp528-529; Schick 1990, pp32-33). Instead, budgetary pressures stressed the virtues of efficiency and expenditure restraint. Taking this lead central finance agencies preferred ad hoc evaluation as a convenient way of reverting to their "oldest purpose—controlling expenditure" (Schick 1988, p532). In other words, the demand for program evaluation had shifted again.

This brief survey of the comparative literature has suggested that changes in the nature of demand have determined how program evaluation has been institutionalised within government. Demand analysis has pointed to the alignment of program evaluation with the budgetary responsibilities of central finance agencies as the most significant response to adverse economic conditions and accompanying shifts in political priorities. However, despite their cogency and consistency, the broad trends identified by the comparative studies above have only partial application to the Australian case. In part this is because this research has confined itself to studying industrialised countries of the northern hemisphere. As a matter of course this introduces some degree of context specificity. Perhaps the most problematic aspect of this is the blanket assertion that fiscal-stress-induced changes in the demand for evaluation have encouraged the adoption by governments of "tactics that are *finance* rather than *policy* led" (Gray & Jenkins 1989, p24 emphasis in original; Pollitt 1993, pp356-359). The implication is that concerns with economy and efficiency necessarily displaced attendant concerns with the effectiveness of policy, and that central program evaluation was shaped by this orientation.

Within these trend lines Australia is somewhat of an errant case. Certainly, in common with other industrialised countries, Australia confronted protracted economic problems during the period of the second wave that compelled governments to adopt similar overall policy objectives. Policy designs for evaluation, however, diverged. Most prominently, as the chapters in Part Two will demonstrate, Australian governments did introduce a formal policy of systematic program evaluation and, contrary to several of the comparative assessments (Schick 1990, p33; Gray & Jenkins 1993, pp189-190), deliberately sought to establish a tight link between resource budgeting and program evaluation. Just as incongruously, the Australian government explicitly directed budget related program evaluation at the effectiveness and relevance of government policy rather than simply cost efficiency.[1] In part this reflects Australia's relatively late entry into program evaluation—at the tail end of the second wave—but it is also a consequence of the gravity of the problems facing Australian policy makers, as well as the distinctiveness of the responses they developed. This applies to the way in which central government was restructured as much as the pragmatic version of managerialism which emerged in Australia as a response to economic stress (cf Halligan 1994, pp4, 11-12). This pragmatism was led by the social justice concerns of the incumbent Labor Party. In what follows we continue this line of analysis by surveying how and why central government has been reshaped in Australia and the way in which central agencies both contributed to and benefited from this process.

Economic Crisis And The Reshaping Of Central Governmnt In Australia

The comparative literature surveyed above pointed to the significance of a number of policy climate factors—political regime shifts, general economic conditions, and the degree of bureaucratic cooperation—in creating a "corridor of action" generating government demand for program evaluation. Given the somewhat capricious nature of public policy climates, it is likely such corridors

---

[1] On both counts the program evaluation policy adopted by the federal government of Canada between 1977 and 1993 warrants inclusion with the 'errant' Australian case. The position taken here is that, although the Australian policy borrowed much from Canadian developments, the Canadian case can be excluded for two reasons. First, the policy was less a response to acute economic stress and more a political reaction to stinging accusations of inadequate internal financial management (Jordon & Sutherland 1979; Segsworth 1990). Second, the policy's links with the central budgetary process were, at best, tenuous (Segsworth 1993, pp99-100; Interviews 17, 28).

are short-lived and that aspiring public sector reformers must take full advantage of the openings when they appear.

[2] This is not to say that reformers are everywhere prevented from *engineering* change, only that the chances of success are maximised where they are presented with adequate policy opportunities (cf Hood 1995, pp105-107).

This section will chart the policy climate prevailing in Australia during the second half of the 1980s and demonstrate how the confluence of a number of climatic factors created a distinct policy corridor that induced demand for, and facilitated the development of, a policy of systematic program evaluation within Australian federal government. Contending that the response to the 1980s economic crisis in Australia was distinctive, the case made comprises three points. First, that this response was marked by economic restructuring that was tempered by the self-imposed policy constraints of the ruling ALP. Second, that this internal economic adjustment profoundly shaped the way in which central government was restructured. Third, that the central economic agencies in Australia had a crucial role in this process that ultimately determined the complexion of program evaluation policy. In sum, political regime change, acute economic stress and strong bureaucratic engagement ensured that a policy corridor appeared in the late 1980s. To begin, let us first contextualise Australian economic difficulties.

Entering the 1980s many of the western world's industrialised countries confronted protracted macro-economic difficulties. The post-war period had delivered enormous increases in world trade that, in the main, sustained continuously rising levels of economic growth and an expansion of the scope and levels of government expenditure. This upward trend of prosperity petered out towards the end of the 1970s as the international economy was buffeted by a number of price shocks, most prominently the OPEC oil crisis of the early and late 1970s. As a consequence a number of seemingly intractable policy problems emerged as common to most domestic economies: an appreciable decline in the rate of economic growth, a marked increase in the rate of both wage and price inflation, and persistently high levels of unemployment. In the face of these difficulties the post-war Keynesian consensus that subscribed to government intervention as the way of smoothing the bumps of volatile economic cycles

[2] Alternatively, it has been argued that public sector restructuring is, in part, a reaction to the onset of economic and political "globalisation"—that is, the new financial, trade and communications linkages that are increasingly integrating the world economy and purportedly diminishing the capacity of nation states to affect policy outcomes (Savoie 1995; cf Hood 1994, pp133-138). The impact of "globalisation" is, however, equivocal. Most obviously, does public sector restructuring represent attempts by the nation state to resist global forces or accommodate them?

broke down (Wright 1981). Its place was quickly filled by the advance of a 'neo-liberalist' political agenda that identified the drag of big government and excessive spending on social welfare as the main impediments to economic performance. Economic crisis had provoked a reappraisal of the appropriate role of the state in economy and society.

Introspection of this type was, however, more pronounced in some countries than others. Castles (1993) contends that the cluster of English-speaking nations experienced the most intense pressure for fundamental policy change. Compared to other western industrialised countries, these nations had recorded both exceptionally low economic growth rates and debilitating outcomes in terms of unemployment and inflation (Castles 1993, pp17-22). It was this mediocre policy record, rather than the size of government *per se*, that undermined broad popular support for the existing policy regimes and catalysed policy reappraisal in the 1980s. Accordingly, the policy responses emanating from each of these countries were a "reaction to the peculiar character of existing policy strategies" that had produced the poor policy outcomes (Castles 1993, p11). Two types of policy transformation emerged within the English-speaking nations—an ideologically driven and concerted retreat from the social welfare state that characterised policy in the United States and the United Kingdom (mentioned in the first section above), and the more pragmatic disengagement with the economic and social regulation that marked the "domestic defence" policy regimes of Australia, Canada and New Zealand. The focus here will narrow to the distinctiveness of the Australian case.

Politics in Australia during the 1980s was clearly driven by economics. With this in mind, a sketch of the economic situation during this period will help set the scene. Certainly, from the very outset of its first term, the Labor Government had to cope with all the most common economic stresses, but these were, in large measure, exacerbated by a number of underlying infirmities in the structure of the Australian economy. These were confirmed by the external balance of payments crisis of 1985-86 (Emy & Hughes 1988, pp2-20; Gruen & Gratten 1993, pp101-102; Kelly 1994, pp196-199, 202-227). Escalating current account deficits could be traced back to the Australian economy's heavy dependence on the export of primary commodities and the declining prices these goods received on world markets. High deficits on the current account required increased borrowing which, in turn, caused ballooning foreign debt levels. The debt problem threatened depreciations in the Australian currency and further aggravation of the current account deficit. There was general agreement that without some kind of policy intervention this cycle might lead to economic atrophy as both economic growth and accepted standards of living were emasculated by debt repayments (Emy & Hughes 1988, pp6-7).

The policy response to this kind of economic stress was, according to a number of commentators, uniquely conditioned by Australian historical development. The antecedents of the contemporary economic crisis could, it was argued, be traced back to the stultifying effects of the system of protection that was established under what has been variously termed the “Australian Settlement” (Kelly 1994, pp1-16) and the policy regime of “domestic defence” (Castles 1988, pp91-104; 1989, pp30-44; 1993). As a small trading economy dependent on the export of primary commodities, Australia had, since its inception, consciously pursued a policy of domestic protection—comprising high levels of tariff protection for industry, centralised wage arbitration framed around need rather than productivity, restrictive immigration polices and a propensity for state paternalism—with the sole intention of insulating the economy from the competitive pressures of the international economy. The end result was that comparatively high levels of regulation had corroded the competitiveness of the Australian economy. The preferred policy response from the mid 1980s was to rectify structural weaknesses in the Australian economy via internal adjustment—broadening the export base to include more manufacturing commodities and extensive deregulation of manufacturing and financial industries so as to open up the economy to international competition and generally improve efficiency (Easton & Gerritsen 1996; Kelly 1994).

Economic structural adjustment emerged as the most significant component of Australian government responses to economic crisis. As Castles has pointed out, this departed from the more explicitly ideological push in the United States and the United Kingdom that sought, first and foremost, to reduce the role and absolute size of the state (see also Self 1993; Savoie 1994). In those countries big government was the perceived problem, not economic competitiveness. These contrasting frames of reference are important precisely because they help determine the distinctive tenor of the Australian response. Structural adjustment was the preferred policy instrument in Australia. However, as it applied to the reorganisation of central government, this strategy was guided by two principal agents—the imperative of tempered economic restructuring advanced by the Labor Government and the opportunistic drive provided by central economic agencies to overhaul the machinery of government. While this is not an atypical interpretation of economic reforms, it does seek to show that the complexion of administrative reform is more problematic than suggested in other “economic rationalist” explanations (Pusey 1991).

The most fundamental policy constraint facing the ALP was accommodating pressures from its political constituency. The economic crisis presented the Labor Government with the dilemma of pursuing wide-ranging economic restructuring at the possible expense of eroding its primary support base, the trade unions (Castles 1993, pp8-9). The dilemma hinged on the ability of the Labor Government to secure economic reform within the broad social justice objectives to which the ALP, as a left of centre party, had traditionally adhered. To some extent, this predicament had been addressed from the outset of the Labor Government in March 1983, when the pre-negotiated prices and incomes accord with the trade union movement, which sought to hold down inflation through wage restraint, came into effect. Although the accord became increasingly more difficult to enforce after the balance of payments crisis in 1985-86, it did establish a distinct consultative approach to economic reform which effectively modified or diluted the effects of economic deregulation—the rationale being to increase economic growth, but at the same time redistribute the economic and social costs of structural adjustment away from the most disadvantaged. Some evidence for this assertion is to be found in the enhancement of social welfare programs during a long period of fiscal contraction in the late 1980s—fiscal priorities shifted overwhelmingly to social equity concerns (Easton & Gerritsen 1996, pp44-45; Campbell & Halligan 1992, pp134-135; although note Swank 1992, pp429-430).

Having said this, there remains, of course, some doubt about the extent to which the accord resonated with the policy traditions of the Labor Party. For a number of commentators, the accord was a veil for flagrant policy reversal (Beilharz 1994, pp126-136; Battin 1993, pp233-239). Under this interpretation, instead of mollifying economic restructuring the accord legitimated a betrayal of the ‘social democratic’ traditions of the ALP that had emerged with the Whitlam Government of the early 1970s. At the same time, others have traced within the relations between the accord and deregulation a continuation of the ‘labourist’ strand in Labor policy that sought to manage the economy more fairly for the benefit of workers (Singleton 1990, pp192-201; Manning 1992, pp13-14, 23-24; Castles, Gerritsen & Vowles 1996, pp10-11). This is not the place to enter the debate on ‘discontinuity’ in Labor Party ideology. The position taken here is that the continuity argument is more plausible, if only because interpreting market liberalisation policies of the 1980s as a “historic departure” ignores other and just as influential—market-oriented—traditions of the Australian state (Bell & Head 1994, pp7-11). In effect, the Labor Government’s response to economic crisis was shaped primarily by its political obligations to “labourism”.

For our purposes it is sufficient to note from this that the major consequence of the Labor Government’s policy of moderated structural adjustment, was that the neo-conservative assault on the size and role of government was deflected towards the restructuring of central government as one way to contribute to increased competitiveness. Economic deregulation was a component of “sound economic management” and any change in the function of the state was projected as an integral part of economic structural adjustment rather than simply deliberate emasculation of government for its own sake (Castles 1993, p9; cf Hood 1995, p107). Labor’s concern for managed economic reform ensured that the restructuring of government would, in some measure, coincide with its social justice concerns. The argument advanced here is that central bureaucratic engagement with government reformers *did* run along this grain and this is precisely why Australian reforms can be marked as distinctive.

One recent interpretation lends some support to this argument. In a comparative study of state reform in small countries governed by left-of-centre parties, Schwartz (1994a, 1994b) explains new managerialism and the restructuring of central government as the product of a combination of political regime change and strong bureaucratic support from central economic departments. Schwartz argues that the economic crisis of the mid-1980s, characterised by sharply rising fiscal deficits, spurred welfare state reorganisation framed around “market mechanisms” and the enhancement of efficiency. The key proponents of this strategy, he argues, were “fiscal bureaucrats” who, in coalition with both new governments and vulnerable industries exposed to the international economy, played a “crucial enabling role” in pushing not for a diminution of the welfare state but a *reorganisation* of the state so as to subject the public sector to efficiency generating competitive pressures (Schwartz 1994a, pp50-53; Schwartz 1994b, pp535-537, 548). Certainly, politicians, particularly in the Australian case, had proposed the broad changes, but “it was fiscal bureaucrats who provided most of the programmatic ideas”. Furthermore, the more inclined central agency officials were to provide detailed plans, the more advanced the reforms (Schwartz 1994a, p53). Australian central agents were judged by Schwartz to be particularly active, so that the contours of increased competitiveness hastened the flow of restructuring in government.

Schwartz’s principal contention is that in driving this government reorganisation central economic agencies pursued their own agenda. Knowingly or not, fiscal bureaucrats conformed to public choice prescriptions in advising governments that rent seeking behaviour on the part of welfare interest groups and public sector employees was the primary cause of uncontrollable growth in government spending (and hence rising fiscal deficits). As these demands increased, the capacity of central agencies to control spending levels decreased. In order to reverse this trend, central economic departments seized on government reorganisation as one way of preserving their traditional “guardian” functions in the public expenditure process (Schwartz 1994a, p53). Reorganisation ran along two lines. The first centralised control over both the volume of public spending and strategic planning with central economic departments (cf Zifcak 1994, p167; Campbell & Halligan 1992). Within these limits, the second sought to control behaviour within the public sector by reconfiguring managerial incentives; authority was decentralised while control was tightened through a focus on the outputs and outcomes of government policies, and greater use of user-pays schemes for government services. On this interpretation restructuring in Australia served two purposes—it assisted Labor Government policies of increasing both general economic competitiveness and fiscal rectitude and, at the same time, conveniently secured the position of central economic agencies within the new structures.

The coincidence of the dual purposes of restructuring explains a great deal about the reconfiguration of central government. Following the balance of payments crisis of the mid 1980s immediate attention focused on fiscal restraint. Economic imperatives dominated, but these were tempered by the policy posture adopted by the Labor Government that angled policy making towards the effectiveness and continued relevance of government policy. Just as social justice was a central concern of economic restructuring, so was it a pronounced characteristic of managerialism in the Australian case, especially when compared with overseas reforms. As a number of commentators have pointed out, the managerial reforms pursued at the federal level in Australia were more comprehensive in their scope and, reflecting the consultative style of the first three Labor Governments, readily embraced a concern for industrial democracy and social equity (Zifcak 1994, pp26, 156; Campbell & Halligan 1992, pp134-135). It is, however, likely that the anti-managerialist critique of the late 1980s—that accused the Labor Government and central agencies of neglecting equity concerns—accelerated the incorporation of these values as part of the new incentive structures framed around "managing for results". These explicitly sought to integrate social justice objectives (comprising equity, equality, access and participation) across all managerial processes (DoF & DPMC 1989, pp1-5; Keating 1989, 1990). This focus, relative to international managerial reform experience, was unique.

As an explanation of the propulsion for central government restructuring, Schwartz's self-interest argument is certainly plausible. It does have theoretical support from recent "bureau shaping" analyses of bureaucratic behaviour in Westminster derived political systems (Dunleavy 1991). These presume that senior bureaucrats are predisposed to seek high status analytical work in central coordinating agencies, distanced from more mundane service delivery. Under these circumstances, there is nothing inimical to their interests in pursuing strategies directed at either fundamental restructuring or cutback management as long as the role and status of central agencies remains intact (Dunleavy 1991, pp200-209, 227-240). In effect, the interests of senior—and particularly central agency—bureaucrats are advanced by cooperating with "politically inspired change" and hence "acting like a well-paid hatchet man [sic]" (Self 1993, pp164-165). Schwartz's interpretation of Australian central agency behaviour sits comfortably with this portrayal.

There are, however, reservations about the capacity of bureau shaping models to explain patterns of reorganisation in government departments. In one study of agency restructuring at the federal level in Australia during the 1980s, Dollery & Hamburger clearly attribute change as much to the influence of strong political direction as bureau-shaping officials (1996, pp493-495). This also points to the first of several problems with the Schwartz thesis. In the first place, it tends to diminish the directive part played by politicians in sweeping administrative reform. Sure enough, in the Australian case Schwartz acknowledges how political regime change provided broad reform parameters, but this comes at the expense of obscuring one of the chief aims of managerial reform—to increase political (ministerial or cabinet) control of the policy agenda as well as making the bureaucracy more responsive to political direction (Zifcak 1994, pp19-21, 154-156). This is an objective that has been common to public sector management reforms in the English-speaking nations (Savoie 1994, pp9-12; Aucoin 1990).

In Australia, efforts at reasserting political control have produced an executive branch regime type that has been termed “political management” (Halligan & Power 1992). In line with the broader economic crisis analysis undertaken earlier, the emergence of political management has been closely associated with the strength of both the political executive *vis a vis* the bureaucracy and the prevailing social and economic climates. Hence, according to Halligan & Power, the current complexion of the executive regime type can be attributed to a fundamental shift in the dominant paradigm guiding Australian public policy—that is, a shift from protected pluralism to strategic central direction in response to persistent economic constraints (Halligan & Power 1992, pp14-16; cf Halligan & Wettenhall 1990, pp32-37). This has been reflected in changes to administrative structures that sought to tighten political control through the agencies of central government by transforming their roles as servicing explicitly political (or policy) needs rather than simply standardising administrative practices. To this extent political management also represented an attempt to modernise central government operations, an agenda pursued under the aegis of consistently strong reformist Labor Governments at both the federal and state levels (Boston & Uhr 1996, pp48-49, 61; Davis 1995; Painter 1987; Sharman & Stuart 1982).

From the preceding discussion of Schwartz’s self-interest thesis of state reorganisation, we can see that a common presumption is that bureaucrats exercise considerable influence over policy making (the most prominent source being, of course, the various branches of the public choice literature). However, whether Australian fiscal bureaucrats *consciously* pursued both public choice diagnoses—that rent seeking behaviour caused unsustainable growth in public spending—*and* remedies is not the important issue here. Relative to the theoretical sophistication of remedies adopted by their New Zealand counterparts, Australian central agents patently did not (Boston 1991; cf Hood 1990, p210). However, there is little doubt that public choice informed the agenda of the Finance Minister for most of this period, Senator Peter Walsh, whose memoirs reveal a remarkable preoccupation with expenditure control and the constant struggle to outwit “rent seekers” (Walsh 1995). The question that remains is, who influenced whom?

Commentators such as Pusey (1991) have laboured this point. Despite enhanced ministerial control, the “economic rationalists” who purportedly dominated the senior ranks of key central agencies unduly influenced the economic and social policy agenda of the 1980s. What we can take from Pusey is the undeniable fact that forceful politicians and equally committed bureaucrats combined to pursue a common agenda of economic and administrative restructuring (Pusey 1991, pp7-9, 64-67, 193; see also Aucoin & Bakvis 1993, p397). We are probably best served by treating this as a convergence of policy positions and that the various processes of restructuring were mutually reinforcing because they intentionally promoted central bureaucratic involvement in policy making. In addition, an interpretation of this type also adds weight to the reforms as “modernisation” since, in seeking to increase bureaucratic responsiveness, the Labor Government was gradually moving towards the kind of politico-administrative “collegiality” observed as the hallmark of central government in Canada (Campbell & Szablowski 1979).

Of course economic disorientation often spurs other problems for governments. One last implication of Schwartz’s thesis revolves around the character of the public choice dilemma facing central agencies—is the loss of control a product of escalating rent seeking or the consequence of increased policy complexity in drastically altered economic conditions? In other words, did Australian governments in the 1980s confront the protracted problems of “administrative overload”, or the “*mismatch* between the demands placed upon government and its ability to *respond* effectively to these demands” (Self 1984, p14, emphasis added)? Given that the higher the level of overload, the larger the required expenditure of effort to coordinate both policy and resources, it would follow that politicians and bureaucrats were actively seeking to modernise the coordinating capacities of central government. As the generation of complex policy problems accelerated, the nature of the central coordination problem was also subject to transformation. In his examination of central agencies in the Australian states, Painter (1987) found that the growth of government in the 1970s and 1980s had precipitated an explosion in both “policy complexity”—the difficulties arising from overlaps and conflicts between different purposes—and “administrative complexity”—the difficulties arising from growing interdependencies among agencies. As both types of complexity increased "problems of coordination undergo qualitative changes" (Painter 1987, pp13-14).

Given the flow-on effect that state developments often had for the federal arena, Painter's identification of a common trend in state administrations of a shift towards a "new agenda of coordination problems with their increased policy content" is particularly cogent (Painter 1987, p152). Policy and administrative complexity were a direct consequence of increased state intervention which had seen previously compartmentalised administrative structures penetrated by "new social and political conflicts". The internalising of complexity compelled political executives and central agencies to devise new sets of procedures that "helped to restructure relationships in a way more relevant to these problems" (Painter 1987, pp74, 152). Techniques such as corporate planning, program budgeting and program evaluation were gradually introduced. Their combined effect was twofold. First, they served to centralise questions about conflicting policy priorities. Second, and more significantly, the techniques served to transform the budgetary process from a preoccupation with accounting for expenditure to a genuine "policy instrument" more attuned to government objectives (Painter 1987, pp102-108). Central agency review intensified as they were brought increasingly into questions of spending department *policy* rather than simply administrative competence (Painter 1987, pp192-193). The position taken here is that the same kind of central government 'modernisation' can be observed at the federal level.

More likely than not, the three implications just discussed add to rather than detract from the value of Schwartz's self interest thesis of central government restructuring. During the 1980s public sector reform initiatives were based on a symbiotic relationship between politicians and central agencies, each pursuing their own instinctive interests for institutional survival. The distinctive restructuring of central government in Australia then was guided by two impulses. The first of these was, most obviously, the Labor Government's policy imperatives of economic structural adjustment tempered by social justice obligations. The second was the drive of central economic agencies, who viewed the restructuring process as an opportunity to both secure their position within new managerial structures and "modernise" the central machinery of government in order to cope with increasingly complex economic and social policy environments.

CONCLUSION

This chapter has been an exercise in orientation. It has provided a preliminary sketch of comparative explanations for the institutionalisation of program evaluation within western industrialised nations that have focused on policy climatic factors. A number of trends were traced which suggested that changes in the nature of the demand for evaluation proved to be particularly significant—program evaluation became the servant of expenditure control. However, these explanations held only partial applicability to developments in Australian federal government where a policy corridor framed by propitious economic, political and bureaucratic elements placed a premium on the restructuring of central government as one way of improving economic competitiveness and efficiency. Although structural adjustment was certainly conditioned by the political imperatives of the incumbent Labor Government, of these three elements central economic agencies played a pivotal role in shaping the new government structures.

In sum this chapter has made a case for narrowing our focus on central agencies—that the institutionalisation of evaluation in Australian government is a function of demand changes that have, in turn, been driven largely by central agencies during a period of economic and state restructuring. The issues that remain are twofold. How have central agencies applied program evaluation within the policy processes of Australian central government? What have been the outcomes? The purpose of the chapter that follows is to construct a framework for analysing these issues.

Chapter Three

Policy Management And Program Evaluation

The purpose of this chapter is to construct an analytical framework for assessing the impact of program evaluation on the work of central agencies in Australian federal government. It contends that program evaluation policies are an important component of an international trend to strengthen the management capacities of central agencies. As such the chapter will explore the nature of the core functions of central agencies within the policy processes of government, an area of the public policy literature that has been characterised by considerable analytical contestability. The argument is twofold. First, that the core functions of central agencies can most profitably be framed around the management of the coordination problem in central government—here referred to as *policy management*. Second, that these policy management capacities have been enhanced in response to the transformation of the coordination problem from an initial concern with the process of government to a sharpened focus on substantive policy concerns. Program evaluation has been a crucial instrument for enhancing these capabilities.

The framework developed in this chapter recognises that program evaluation's contribution to policy management capabilities has two aspects. The first is substantive and refers to the response to the heightened complexity of coordination problems associated with policy and resource coordination. The second is procedural and deals with the institutional constraints which ensure that policy management can only be achieved within the established 'rules of the game' of bureaucratic politics determined by central agencies. In order to examine the contribution of program evaluation to policy management this chapter will concentrate on two key questions. Has program evaluation, by altering the rules of the game, recast the way that the policy management function is approached? Does program evaluation lead to the over-extension of central agency policy and resource coordination capabilities (a condition that can be termed policy overstretch) by forcing these agencies increasingly to assume both procedural *and* policy competencies?

The chapter proceeds in three steps. First, it briefly outlines a rationale for the study of central agencies as distinct and influential political actors within the policy making process. Second, it will develop the policy management function of central agencies as essentially a response to intensified problems of coordination in government by examining the various roles of central agencies. Third, it concludes with an examination of program evaluation as an augmentation of policy management capacities and derives the two key questions for this thesis.

Central Agencies As Political Actors

A recent impetus for the study of central agencies has been the emergence of a new focus on departments as distinct political actors. This stance proceeds from a recognition of the limitations of much of the extant literature that analyses the exercise of power within central government confined to the “artificial bounds” of the constitutional approach of cabinet government. As some authors contend, a major shortcoming of this approach is that it imposes a “normative ideal, a constitutional theory” that leaves us with an incomplete if not distorted representation of political reality (Dunleavy & Rhodes 1990, pp3-5; Smith, Marsh & Richards 1993, pp569-571). The presumption of cabinet primacy—or, alternatively, prime ministerial power—has precluded proper analysis of the “core executive” that, according to its protagonists, can be taken:

> functionally to include all those organisations and structures which primarily serve to pull together and integrate central government policies, or act as final arbiters within the executive of conflicts between different elements of the government machine (Dunleavy & Rhodes 1990, p4).

For these critics, the record on core executive studies is disappointing. Much of the structure, particularly in terms of the role of departments, remains “virtually unspecified” (Dunleavy & Rhodes 1990, p21). As a more recent survey has lamented:

> Little [of the literature] has come to terms with the position of departments in the policy process. Departments are the key policy makers. Most policy is made within government departments but there is very little research on the way in which departments make policy and how their structures and ideologies affect policy outcomes (Smith, Marsh & Richards 1993, p589).

Of these organizations and departments, central agencies are the most influential and, by and large, least examined. What is now required is an “accumulation of systematic organizational studies of central department’s structures, values and decision making behaviour” (Dunleavy & Rhodes 1990, pp20-21).

Although there is considerable merit in this research agenda—particularly its renewed focus on the pivotal policy role of government departments—it can only be taken so far for two reasons. First, as we discussed in the preceding chapter, a significant spur to the redesign of central administrative agencies over the past decade has been the externally imposed imperative of reinforcing the policy making superiority of cabinet (Dunleavy & Rhodes 1990, pp15-16; Aucoin 1990; Savoie 1994). Therefore, it makes little sense to bring departments to the foreground by pushing cabinet (or chief executives) to the background. Certainly, some commentators contend that public sector reform during the 1980s was marked by an inadvertent return to the conventional politics-administration dichotomy, since politicians held a jaundiced view of the power of bureaucrats that manifested itself in managerial reform aimed squarely at curbing the role of public officials and central agencies in policy making, and replacing them with partisan policy advisers (Savoie 1990, pp12-23, 341-345). However, given that in a number of countries central agencies were seen patently to have driven managerial reforms (see Chapter Two), the general applicability of this assessment rapidly diminishes. Rather, the crucial policy-servicing role played by officials in central agencies means they remain highly influential. Second, contrary to the assertions of proponents of core executive studies, there is a considerable literature that has concerned itself with the policy role of central agencies within cabinet governments.

As a starting point for examining the role of central agencies we should provide some orientation by defining the term “central agency”. It should be noted, however, that the brief discussion that follows serves only as a precursor of more detailed analysis of central agency functions in the context of policy management. This appears later in this chapter.

One noted scholar of public administration has suggested that the meaning of the term "central agency" remains essentially contested (Boston 1989, p6). There are, indeed, are number of definitions that place different emphases on different structural characteristics. Boston himself adopts the following definition—"central agencies are government departments or ministries with broad control, review, co-ordinating, and policy vetting functions" (Boston, 1989, p6). Functional definitions in this mould abound (Campbell & Szablowski 1979, pp2-3; Campbell 1983, pp17-18; Campbell & Halligan 1992, p50). As this indicates central agencies are defined primarily by reference to the type of functions they perform. While the above definition outlines some of these functions, taken in isolation it fails to tell us what it is that make central agencies *distinct* from other administrative structures. Campbell and Halligan go some way further. They maintain that a central agency can be defined by "two primary characteristics"—the absence of line (or spending) department responsibilities and the overriding importance of "systemic roles or responsibilities which extend across line departments" (Campbell & Halligan 1992, p43). Central agencies then have structures, roles and an "ethos" which distinguishes them from other agencies. For the purposes of examining central agencies and the coordination problem in government, the following general definition will be adopted:

> a central agency stands above other departments in that it performs functions which are thought to be crucial to the *common* interests of government departments, and which relate to matters of major importance (Campbell & Szablowski 1979, p2).

CENTRAL AGENCIES AND COORDINATION

If central agencies are responsible for managing the "common interests" of government, then it follows that perhaps their most fundamental task is to coordinate and arbitrate what are usually disparate departmental interests. The purpose of this section is to describe the nature of the coordination problem in central government. This is an area that has been subject to considerable debate. In public administration, coordination is the most commonly accepted way of dealing with the range and complexity of government activities, where policy objectives must be divided up for ease of consideration and manageability. At its base, coordination is about reconfiguring these diverse purposes so that they are aligned with current objectives (Peres 1978, p249; Spann 1979, p411). Hence, according to Painter, coordination is “a problem arising as a secondary matter out of the prior need to subdivide and specialise” (1981, p266).

It is, of course, a commonplace that it is better that one’s activity be coordinated than incoherent and out of control. In politics it is an imperative. Davis (1995) has convincingly demonstrated why governments in Australia—and parliamentary systems in general—consistently strive to attain coordination in both policy and administration. The adversarial nature of the parliamentary system, whereby ministers are responsible to parliament for the stewardship provided by government departments, means that:

> [p]olitical survival requires a government to appear coherent and united, in control and able to account for the resources in its care (Davis 1995, pp16-17).

Above all else, coordination, or at least the appearance of coordination, dominates considerations of political rationality. Given the resumption of political control over the bureaucratic apparatus that has characterised the waxing of central agencies, there is reason to believe that coordination retains its title as “an undisputed virtue” in government (Davis 1995, pp16, 141; cf p137).

As an administrative platitude, coordination has few detractors. However, in government, coordination is essentially a political process that is subject to a number of fiercely competing definitions. Most commonly, and especially by those who are being ‘coordinated’, coordination is seen as either central control or simply coercion—indeed, coordination as hierarchical control has been categorised by one study as the “benchmark model” (Minnery 1988, p259). Peres, for example, equates coordination unequivocally with power:

> Coordination is one of the fraudulent words of politics and administration. It dresses neutrally to disguise what is nakedly pure political form. Coordination is a political process by which the coordinated are made to change their value positions, their policy conceptions and their behaviour to conform with the conceptions of the coordinator. Force or persuasion can give rise to the change (1974, pp151-152).

Peres focuses on what he terms the “capacity” of governments to coordinate. Since the power available to governments is scarce, and because coordination is a function of power, it follows there is a finite capacity in administrative systems to coordinate (Peres 1978, pp252-253). This capacity is sensitive to changing organizational and policy conditions and "must be used judiciously and conserved" (Peres 1978, p253). Hence, adopting techniques to increase coordination will necessarily require a trade off in government between policy performance and coordinating capacity. In other words, either government performance will be diminished as coordination is imposed, or coordination must be neglected in order to improve the policy effectiveness of government.

Coordination as power is, in Peres' view, a zero sum political game. As a form of control, it is probably more accurate to view coordination capacity as occurring in degrees—it is not an all or nothing affair (Spann 1979, p246). If coordination capacities at the centre are sensitive then they are also, to some extent, self-limiting, since instruments for coordination can exacerbate the coordination problem through the costs it can impose. As Spann notes, the continuing relations between coordinators (central agencies) and the coordinated (spending departments) is "one of the great sources of conflict in government" (Spann 1979, p420).

This equation of coordination with power has, however, been questioned. Pressman & Wildavsky, for instance, agree that coordination is a "deceptive" administrative good, but argue that this is because it can be seen as *either* coercion or consent (Pressman & Wildavsky 1979, pp133-135; Wildavsky 1979, pp132-133). As a remedy for debilitating fragmentation, they are especially sceptical as to whether coordination holds any prescriptive value for policy makers. This arises because calls for more coordination appear to be tautological in that they imply that the only way to acquire better coordination is already to possess it.

Alternatively, coordination has been divorced from considerations of the exercise of power. One prominent example of this is coordination as partisan mutual adjustment, which sets out to "deflate" central coordination (Lindblom 1965, p293). This is a process of political coordination that contends that individuals can coordinate their actions with each other without anyone's coordinating (that is, controlling) them, and doing so in the absence of a dominant common purpose. In a situation where decision makers are both interdependent and self-interested, coordination can be achieved by adaptation on the part of decision makers—that is, partisan mutual adjustment (Lindblom 1965, pp28-32, 154). The important characteristic of mutual adjustment is that every decision is part of the process of coordination and hence self-validating. It is precisely this attribute that sees Westminster-style systems of government leave very little room for coordination between departments by partisan mutual adjustment, since coordination capacity, conceptually at least, is located at one place—cabinet—rather than at crucial decision points in the executive structure (Painter & Carey 1979, pp77-78).

All of the above conceptions of coordination view it primarily as “a political value pursued by procedural means” (Davis 1995, p18). Another way of framing coordination is to treat it as a unique procedural value pursued by central agencies. It is in this vein that Painter develops the notion of the coordination principle in public administration (Painter 1981; Painter 1987). This contends that coordination is a distinct procedural value within government that is concerned primarily with managing conflict (Painter 1981, p276; Painter 1987, p9). In determining central agency functions, the principal focus should be on procedural values—or the ‘rules of the game’ that central agencies “promote and defend”—rather than policy values that are “the subject of contemporary political debate (Painter 1981, pp273-274; cf Davis 1995, pp19-20). Crucial to his analysis is a further distinction between policy coordination as both an objective and a process. As an objective, policy coordination means striving for a consistent set of policy outcomes—an aim Painter dismisses as a “chimera in public policy” (1981, p275)—while as process, policy coordination is a policy making contest in which overlaps and inconsistencies between different subject matters are continuously addressed.

The proper role for central agencies, as Painter sees it, is to create the conditions under which the different forms of coordination can be achieved across diverse policy sectors. In this way he explicitly rejects coordination as central control, preferring instead:

> coordination *among* the parts rather than *of* the parts by some controlling body or person. There is no necessary connotation of a higher overriding purpose either. Rather, there is a process of resolving differences, with the autonomy of the parts being stressed rather than their subordination (Painter 1987, p8, emphasis in original).

Here coordination is essentially a process of bargaining and accommodation to conflicting interests, since there should be “no necessary connotation of a higher overriding purpose”. While this emphasises the importance of central agency coordination as a distinct management role, it also dilutes somewhat their role in central guidance.

Other interpretations lend credence to Painter's analysis, while at the same time imputing an explicit guidance function to central agencies. The crucial role played by procedural values, which Painter stresses, is also the focus of some cybernetic perspectives of control in government (Dunsire 1990, 1991; Hood 1991). Cybernetics is the science of guidance, evaluation and control as they apply to the dynamics of natural and artificial systems. Systems are observed to be self-controlling because they consist of many "servo-mechanisms" that monitor and regulate activity according to set limits that are designed to maintain performance. Unlike natural systems, such as the human body which has internally set limits, an artificial system, such as government, needs to "be controlled" (Dunsire 1991, p327). Given the complexity of government operations, central agencies set these limits—and exert control or coordination—via procedural objectives. Just as Painter suggests, procedural principles are the preferred option because they are both public sector wide and enduring. This is in contrast to policy objectives which are "demonstrably weak" because they are a function of political demand and easily reversed (Dunsire 1991, pp335-336). Of course, procedural values can also contradict, but tensions of this type simply give more systemic steering capacity to central agencies by allowing them to select the weighting or priority they give to different procedural objectives (Dunsire 1991, pp338-344; Hood 1991, pp354-358).

A similar approach is adopted by Kaufmann (1991a, 1991b) who is concerned with the "growing entropy in the public sector" which sees the amount of resources needed for coordination growing at a faster rate than the output in services to the public (1991a, p6). This is a direct consequence of escalating policy complexity. In rejecting the notion of a single coordinating mechanism, Kaufmann focuses on the public sector as a highly differentiated set of interorganisational networks focused on specific policy sectors and characterised by a growing interaction between public and private sectors (Kaufmann 1991a, p8). This interdependence generates long chains of actions that need to be linked or coordinated. Again coordination can best be achieved by manipulating the configuration of rules by which actions occur and integrating the crucial functions of guidance, evaluation and control (Kaufmann 1991a, p8; Kaufmann 1991b, p215).

In rejecting the dominance of hierarchical coordination, Kaufmann, Dunsire and Hood argue that there is no optimal mode of coordination; rather the question is how to shape the patterns of coordination to enhance system learning capabilities—effective policy making is promoted by a system which promotes coordination through learning processes. They are concerned with the “steerability” of governments and equate coordination with control, since there must inevitably be a link between evaluation processes and control (Kaufmann 1991b, p222). Hence, while coordination is not necessarily exercised from a central location, the settings for, or configuration of, rules are.

Framing the coordination function of central agencies as procedurally oriented is significant because it provides a base against which to analyse changes in the roles of central agencies. Painter, for one, has discerned a shift in the role of central agencies which has re-oriented the coordination principle in such a way as to give “greater prominence to problems of *policy* coordination” (Painter 1981, p279). As discussed towards the end of Chapter Two, the spur to this changing posture has been increasing administrative and policy complexity combined with political imperatives to reassert control over policy making. The upshot is that central agencies have been forced to increase their involvement in the contest over policy. As Painter points out, this has implications for the role of central agencies since they “tread a fine line” between detached observation of policy disputes and partisan intervention to secure centrally desired objectives. The former can be pursued, but at the expense of any “whole-of-government” perspective, while the latter strategy of intervention means duplicating departmental efforts and running the risk of both increasing conflict and reducing coordination (Painter 1981, pp279-280)

An important manifestation of this development has been the increasing use by central agencies of tools such as program budgeting and program evaluation to tackle the new types of coordination questions (eg Painter 1987, pp91-93, 99-108, 126-130, 152, 183-184). These have “generated at the centre a new set of resources for exercising influence in the coordination process” (Painter 1987, p192).

Observations of this nature raise a number of questions. How have these new sets of resources influenced the coordination processes of central agencies? If central agencies do pursue policy coordination through intervention, what is the likelihood that coordination capacities will be impaired by what may be called “policy overstretch”.

[1] This refers to a situation where the "global" substantive policy competency of a central agency is insufficient to ensure effective coordination. Painter points to this. As the next section will show, most prescriptions for the appropriate role of central agencies advocate finding a balance between intervention and refereeing. In this case Painter suggests the need to keep a tight rein on any movement into the realm of substantive policy since this is likely to diminish attempts to promote the most favourable conditions for processes of policy coordination. In effect the balance may be shifting—perhaps unwittingly—away from the manageable task of ensuring processes for policy coordination towards the considerably more difficult task of developing (and duplicating) credible government-wide policy analysis capabilities.

THE POLICY MANAGEMENT FUNCTION

To this point, the discussion has traced how central agencies in Westminster-type parliamentary systems have developed in response to markedly altered environmental conditions, as well as the accompanying changes in the fundamental coordination questions facing those agencies. The purpose of this penultimate section is to outline the scope of the policy management function of central agencies. The concluding section will introduce program evaluation as an instrument for policy management. It will conclude by summarising the argument of the chapter and enumerating the major questions which later chapters will seek to answer.

In the preceding discussion of coordination, we saw how some students of central agencies viewed their role primarily as the promotion of important

---

[1] The term "policy overstretch" is derived from Kennedy's discussion of "imperial overstretch" in relation to the purported decline of the United States of America as the pre-eminent world economic and military power (Kennedy 1988, pp514-535). Kennedy suggested that entering the 1990s the United States had to confront:

> the awkward and enduring fact that the sum total of the United States' global interests and obligations is nowadays far larger than the country's power to defend them all simultaneously (Kennedy 1988, p515).

In an analogous manner, "policy overstretch" refers to the difficulties that a central agency faces when it takes on more policy coordination obligations than it can deal with adequately because it can never possess appropriate levels of policy expertise. The best example—and the one to be examined in this thesis—is the central budget agency (the Department of Finance).

procedural values that govern relations between competing policy arenas. An explicitly interventionist policy role was not, until quite recently, countenanced. Other students have sought to waive a strict separation between policy and administrative coordination roles and develop models of the core functions of central agencies.

Boston (1989, pp6-7) outlines a number of “guiding principles” for the organisation of central agencies. The first imperative is to design them with the aim of ensuring good cabinet government through the provision of high quality policy advice. This principle remains undisputed. Supporting frames are less well established. For Boston the desired characteristics include more than one central agency, each with separate responsibilities, and that the principle of “multiple advocacy” be promoted in the provision of policy advice. In his examination of political executives, Campbell (1983, pp17-18) suggested that central agencies “take or share the leading role in one of three broad categories of control and coordination functions”—first, the development of strategic plans for government and the enforcement of compliance; second, the development and integration of economic and fiscal policies; and third, the allocation and management of resources. Campbell & Halligan (1992a, pp50-57) followed broadly this schema in enunciating the following four functions—strategic planning and coordination, advising on economic policy, resource management and managing the public service. According to Campbell & Halligan, managing the public service had been “neither accorded much attention, nor clearly differentiated” in the existing literature and was added to reflect the near universal reform priorities of recent political executives (1992a, pp163-165).

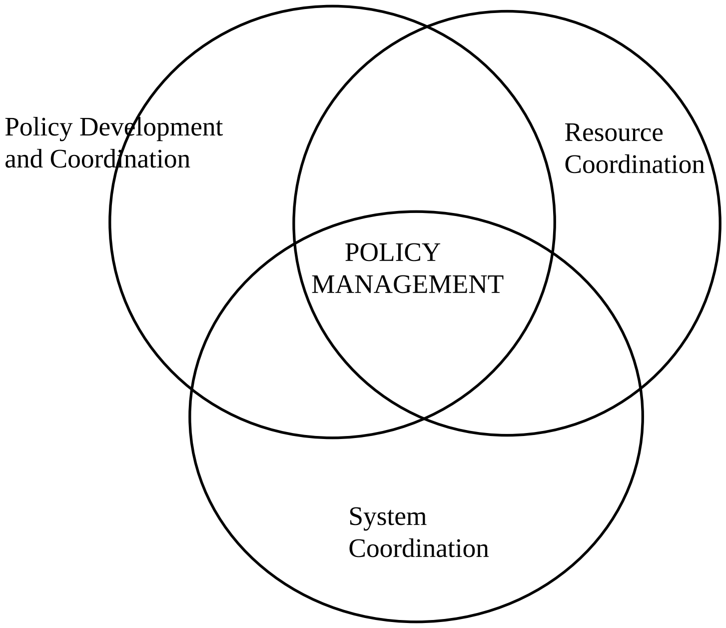
To other commentators expansive functional definitions of this type are cumbersome. Davis (1995, pp138-139) argues for two distinct coordination roles for central agencies—one dealing with policy and the other concerned with the budget and management. The policy domain, as he terms it, centres on cabinet and the procedures for submissions and their prioritorisation. Distinguished from the policy domain is the administrative domain which is coordinated by the finance and public service agencies. Their role is to both establish a detailed framework for management and ensure adequate scrutiny of financial and program performance. With even more economy, Higgins & Borthwick (1990, p45) define the role of central agencies generally as having a “broader policy overview role”, their station being to “represent the wider, general community interest”.

We could continue in this fashion, but these examples convey much of the flavour of this literature. There are considerable overlaps in the functional definitions outlined above, and persisting with them seems only to cloud the focus on the core functions of central agencies. As a consequence, it might be more instructive to develop a functional definition that not only clarifies the core functions mentioned above but that also seeks to capture the respecified policy coordination responsibilities of central agencies. It is submitted that “policy management” demonstrates a fitness for this task. Policy management is a generic term that will be used in this thesis to describe the core functions of central agencies in managing the policy process within contemporary Australian government administration. As **Figure 3.1** below illustrates, the policy management function comprises three dimensions.

The first of these dimensions is *policy development and coordination*. This corresponds with Painter’s discussion of policy coordination. It is composed of two elements. First, at its base, it is concerned “with the substantive interrelations between different subject matters of policy”. To this end, policy coordination revolves around the “resolution of conflicts arising from overlaps, the search for priorities between policies and the overlaying of broader perspectives on narrower sectional views of policies” (Painter 1981, pp275-276). This is roughly policy coordination as process. Second, it is also directed consciously at the development of new policy and the integration of existing policies. In part this resonates with what others have variously termed “strategic planning” (Campbell & Halligan 1992, pp52-53) and “positive coordination” (Self 1980, pp29-30). Here, the central agency takes on the obligation of initiating and guiding the development of new government policy, most commonly policies with either a whole-of-government perspective or that cut across a number of policy sectors.

The policy development and coordination dimension has primarily been the responsibility of those agencies servicing either the chief executive (prime minister) individually or the cabinet collectively. At the federal level in Australia this role is assigned to the Department of the Prime Minister and Cabinet. Increasingly, however, finance agencies have moved into this area as part of their augmented responsibilities in times of fiscal and policy stress.

Figure 3.1 Core Function of Central Agencies: Policy Management and its Dimensions



The second dimension is *resource coordination*. This dimension refers principally to the allocation of resources and the management of the expenditure budget process. In the context of this consideration of policy management, this dimension excludes the economic policy advising function which, depending on whether it is combined in the one agency with resource allocation obligations, is treated as either a separate function (Campbell & Halligan 1992, pp53-55) or an integral part of the same function (Campbell 1983, pp17-18). Resource coordination or budgeting is often seen as the "one form of obligatory coordination within government" (Self 1980, p36; cf Painter 1987, p132). In this context the role of central agencies is twofold—they issue and enforce the procedural rules of the game, while at the same time engaging as a participant in the process. As the next section will demonstrate, this situation places the central budgetary agency, which is responsible for servicing both the finance minister and cabinet, in a precarious position.

The third dimension is *system coordination.* This last dimension covers a wide range of responsibilities directed at maintaining the framework within which government administration operates—that is, maintaining the structural conditions for effective policy management. It consists of two aspects. The first accords with the promulgation and preservation of the procedural framework for government institutions. This includes—but is not confined to—the accountability framework for government agencies, the procedural framework for personnel management and the establishment of standards of appropriate conduct and service delivery. The second aspect is what has been referred to as “change management” or the conscious pursuit of systemic reform (Campbell & Halligan 1992, pp56-57). The political impetus for more responsive bureaucracies has seen the accretion to central agencies of a formal responsibility for continuing administrative reform and the constant improvement of management practices. Although this dimension was primarily the responsibility of autonomous agencies concerned with public service establishment, increasingly it has been dismembered and allocated to a number of central agencies (commonly as part of the reform process itself).

As the core function of central agencies, policy management embraces the three dimensions taken together. Each dimension contributes to policy management as a response to the transformation of the coordination problem in recent years. This is not to say that each central agency in an administrative system must possess all three dimensions of the function. Part of the rationale of this thesis is to examine three observations that can be made about central agency responsibilities. First, that particular central agencies are well suited to roles defined by specific dimensions. Of course, to some degree, agencies will inevitably have responsibilities that encroach into other dimensions but, by and large, these would normally remain secondary. These are represented by the areas created by the overlapping circles in **Figure 3.1.** Second, that it might be deleterious if any one agency, given its original orientation, was to enter into the core competency of another dimension, and hence entertain the possibility of policy overstretch. This is not simply a case of avoiding the friction that can arise from “jurisdictional politics”, but rather one of preventing the overextension of existing policy coordination capacities (Painter 1987, pp191-194). Third, that the techniques used to improve policy management capacities may well be increasing the likelihood of this and hence both expending scarce coordinating capacity and reducing administrative system performance.

Policy management, it is argued, is an appropriate label for the core functions of central agencies since it is singularly concerned with the unavoidable question of coordination in government. The term itself has been invoked in the relevant literature, although somewhat obliquely—Campbell (1983, p23) refers to “policy management responsibilities” as the essence of central agency advisory systems. It accords well with the view of coordination adopted here—as a procedural value which frames the role of central agencies as primarily one of management (Painter 1981, p276). Central agencies are charged with the task of managing conflict within the policy making processes of government. As this chapter has argued, the nature of this task has changed as the coordination questions facing policy makers have changed in response to altered policy environments. At the same time, policy management also resonates with the most recent "paradigm shift" in public administration towards improved management capacities (Hughes 1994, pp1-22).

Evaluation As An Instrument For Policy Management

The focus of this final section is on analysing program evaluation as an instrument for the policy management function of central agencies. Its point of departure is the issue of the institutionalisation of evaluation within the processes of policy management. The principal issue is how evaluation has or has not contributed to the performance of this function. Our interest is to be narrowed down to two of the three identified dimensions of policy management—policy development and coordination, and resource coordination.

There is a sizeable literature that examines what is here described as the policy management function of central agencies. For the most part these have been broad inquiries into the 'rules of the bureaucratic game', especially as it relates to the expenditure budget process. The concentration on the budget has been fuelled, in part, by the paucity of information about both the actual processes used in this highly confidential setting and the respective roles of organisations and individuals. For our purposes we can note that the literature operates at two levels. The first represents a mapping exercise of the framework of central agency power. It charts the various networks of control that specifically relate to the expenditure budget process. The second seeks to analyse the impact of specific instruments of control. Of most interest here is investigation of the introduction of policy analytical techniques to improve budgetary and policy making processes. As we will see, analysis at this second level is remarkably uniform in the sombre conclusions it reaches about the efficacy of these techniques. The two 'levels' of research will now be examined in turn.

At the first level of what can be termed "framework" studies, two in particular stand out. In the first of these, Heclo & Wildavsky (1974) set out to describe the expenditure process as it "actually operated" in British central government. It is the seminal work on central agency behaviour in Westminster-type political systems—the hallmark of the study being the disclosure of how Treasury exercises control in Whitehall through the *informal* networks that exist between departments. Heclo & Wildavsky's focus was on the behaviour of the small number of "political administrators" (senior civil servants and ministers) at the very centre of government (1974, pp2-3). They interpreted the work of political administrators in terms of "community and policy"—community being the personal relationships between politicians and administrators, while policy referred to "governmental action directed toward and affecting some end outside it" (Heclo & Wildavsky 1974, pxv).

By comparing Whitehall to a "village society", Heclo & Wildavsky demonstrated how the Treasury maintained control in a system based on close personal relations and the accepted currency of reputation and trust (1974, pp14-21). The system is sustained by the "circulation pump of career mobility" which ensures that the Treasury ranks are familiar with disparate departmental interests (Heclo & Wildavsky 1974, p82). Within this communal context, control is exercised through essentially informal processes of bureaucratic relations:

> Treasury influence rests not on hard-nosed interpretation of formal powers but in personal networks, sensitive bargaining and up-to-date information that operates to create habits of mind leading to anticipation of Treasury reaction (Heclo & Wildavsky 1974, p380).

As an example, the expenditure budget process is particularly apposite, if only because budget negotiations are characterised by the expectation that it is the bargaining process itself, as distinct from the substantive issue in question, that matters most (Heclo & Wildavsky 1974, p89). Not surprisingly, the authors conclude with some dissatisfaction that in most instances policy improvement remained subordinate to community maintenance (1974, p366).

Weller & Cutt (1976) replicate the "social anthropological" research of Heclo & Wildavsky in explaining the techniques of expenditure control employed by the old Australian Treasury. Starting from the premise that "the procedures of budgeting will often determine the outcome of the process" (Weller & Cutt 1976, p2), they found the same informal networks of control operating and noted that similar notions of 'community' held sway in central Australian government. However, while formal Treasury control was based on the combination of macroeconomic policy management and detailed expenditure control, informal control was based upon "uncertainty"—the uncertainty fostered by the absence of formal procedures for financial control and the imprecision that departments inevitably faced in estimating their costs (Weller & Cutt 1976, pp 34-35, 50-51, 62, 146).

Both studies cast light on two important issues. The first is that the exercise of power by central agencies is not static. No matter how powerful the Treasury or an equivalent finance department is within the bureaucratic community, it is likely to be impotent during a fiscal climate propitious to increased spending and neutered if it is represented in cabinet by a weak minister (Heclo & Wildavsky 1974, pp132-133; Weller & Cutt 1976, pp24-26, 92-98). This, of course, is not to deny the strong influence of central agency officials in the policy process, but simply to assert that there are limits to the reach of politico-administrative “collegiality” (Campbell & Szablowski 1979; Savoie 1990). The second, and arguably more significant conclusion, is the recognition that the influence of formal procedures and techniques is often shaped by the informal networks of bureaucratic relationships. The force of individual personalities, ‘unofficial’ practices and partisan departmental behaviour are all significant determinants of the outcomes of administrative processes since they constitute the accepted rules of the expenditure budget process (cf Savoie 1990, pp70, 87-92). Attempts to augment these rules, most commonly through formal analytical techniques, are often a source of protracted dispute between central agencies and spending departments. A fundamental issue is how these disputes are settled.

The second level of the literature explores this last issue by examining the effects that new budgetary and policy analytical techniques have on the rules of the policy management game. This literature is primarily concerned with the chequered experience of central finance agencies and policy analysis in both the United Kingdom and at the federal level in Canada. In both instances, attempts to bolster the processes supporting cabinet decision making foundered on the rocks of bureaucratic intransigence and ministerial supineness.

The aborted effort to introduce greater analytical capacity to the budgetary process in the United Kingdom during the early 1970s has been the subject of intense scrutiny (Heclo & Wildavsky 1974; Gray & Jenkins 1982, 1985; Campbell 1983). The existing system of expenditure control was based in the public expenditure survey (PES) conducted by the Treasury since the early 1960s. The principal object of the survey was to better inform cabinet decision making by providing both a “collective briefing” projecting forward policy commitments and accepted guidelines for negotiations on expenditure directed specifically at reducing the work flowing up to budget cabinet. Although it was framed as a planning exercise, its application sought to impose spending constraints. However, instead of calling the budgetary tune, PES became an instrument on which spending departments played their own tune. According to Heclo & Wildavsky (1974, pp232-238), PES "enshrined incrementalism to the nth degree" because while it prevented departments from exceeding guidelines it also committed Treasury to maintaining expenditure at projected rates. A focus by the PES on resources rather than the objectives and results of programs was diagnosed as the problem; the solution was to assess programs against objectives and incorporate this information within the PES process (Heclo & Wildavsky 1974, pp271-275; Gray & Jenkins 1982, pp432-435). Thus was born Program Analysis and Review (PAR).

PAR was essentially a program of fundamental reviews of the relevance of government program objectives and their use of resources that corresponded with the PES, the conduct of which was left to individual departments. Introduced at the beginning of the 1970s, PAR failed to see out the decade. The most comprehensive analysis of PAR attributes its demise to an abject failure to satisfy three crucial preconditions (Gray & Jenkins 1982, pp443-449; Gray & Jenkins 1985, pp110-113, 128-135). The first of these were technical conditions—there was, throughout the administration, no accepted methodology for undertaking PARs. This hindered the diffusion of policy analysis and ensured that when undertaken reviews consumed disproportionate amounts of resources. The second set of conditions placed organisational obstacles in the way of reviews. As a systematic tool for increased coordination, PAR required a "coherent, cohesive and strong central organisation" (Gray & Jenkins 1982, p445). Authority for coordinating PARs was fragmented, with both the Treasury and the ill-fated Civil Service Department sharing responsibility. From its inception this division emasculated the supporting role PAR was to play within the PES, even after it was later shunted to the Treasury (Heclo & Wildavsky 1974, pp277-278). The third set of preconditions relate to political congeniality—in sum, PAR failed to receive tangible support from cabinet ministers, the group of decision makers it was ultimately directed at (cf Heclo & Wildavsky 1974, pp296-298; Campbell 1983, p214). In the end PAR spiralled to its demise because it failed to add anything new to the existing budgetary and policy processes.

More recent analyses of the PES have confirmed assessments of this type for other policy evaluation initiatives (Thain & Wright 1992a, 1992b). Certainly, given the constraints imposed by fiscal stress, the survey has shifted even more appreciably from concerns with planning to tighter control of expenditure.

However, the rules of the budgetary game have not shown a corresponding change (Thain & Wright 1992a, pp19-22). Although spending departments have been obliged under the survey to justify the continuance of base line expenditures with quantitative measures of output, Thain & Wright are

> confident that such data is not used to determine relative value for money between programmes, or that in a bilateral negotiation it often proves decisive (Thain & Wright 1992b, pp208-209).

Indeed, at the cabinet level they contend that effectiveness information is either not made available or, if it is, it is “largely disregarded” since the information does not allow ministers to make comparisons between different programs (although this is compounded by a PES process which quarantines bids from collective consideration by cabinet committees and refuses to reopen bids settled at the bilateral stage of negotiations) (Thain & Wright 1992b, pp212-213). Their doleful conclusion is that “there is little evidence that the cabinet or star chamber is much influenced by either economic or financial calculus” (1992b, p223).

Even though the Canadian federal government displayed a good deal more patience with a formal program evaluation policy, the budgetary experience there tells a remarkably similar story of inter-departmental rivalry and political indolence (Campbell 1983; Savoie 1990; Segsworth 1993). A good point of entry is Savoie’s (1990) analysis of central agency power relations which seeks to isolate those aspects of the expenditure budget process in Canada that help explain the alarming growth in government spending over the preceding two decades. The objective of his study is to gauge the power of central agencies (the guardians) in their primary task of restraining public spending (Savoie 1990, pp5-9, 19). Savoie’s principal contention is that despite the augmentation of finance department armouries with ever-increasing policy analytical budgetary techniques, these agencies were ultimately unable to impose financial discipline. His conclusion was that the spenders usually had the upper hand on the guardians (Savoie 1990, pp320, 326-327). In reaching this finding he pointed to a number of factors that constrained central agencies. Some of these are context specific, such as the role of rampant regionalism in Canadian politics that constantly fuelled new spending, while others were more transferable. The latter included the distinct absence of strong political commitment to expenditure reduction and the bureaucratic facts that the “spenders quickly learn to play under any rules” and are “quite capable of circumventing the obstacles of any budget process” (Savoie 1990, pp333-341).

For most of the 1980s the defining landmark on the Canadian expenditure budget process was the policy and expenditure management system (PEMS). Like the expenditure survey in the UK, PEMS was designed to streamline cabinet decision making by locating the responsibility for fiscal discipline directly with ministers (Savoie 1990, pp61-68, 329-330; Van Loon 1983, pp257-261). Under the system, cabinet committees were assigned spending limits or "envelopes" as well as a policy reserve for any new policy proposals. It was expected that if ministers or committees wished to approve spending that exceeded the reserve, they had first to review the effectiveness of programs and find savings from elsewhere in their envelope. In this way, PEMS was supposed to force trade-offs.

Prior to, but in parallel with the development of PEMS came a formal policy of program evaluation that obliged departments to evaluate the efficiency and effectiveness of their policies on a systematic basis (Segsworth 1990; McQueen 1992). The policy was coordinated from a specially created central agency—the Office of the Comptroller-General (OCG). Even though PEMS encouraged policy evaluation in order to inform expenditure policy trade-offs, in retrospect evaluation was accorded a "minor role" in central budgeting (Segsworth 1993, p100). Indeed, by the tail end of the 1980s PEMS had been deemed a failure, for several reasons. First and foremost, as noted earlier, ministers proved extremely reluctant during a period of fiscal stress to volunteer spending cuts. Of equal import, however, was the quality of the program performance information coming from program evaluations. A number of studies have observed that evaluation was not used in the budgetary process precisely because the evaluation priorities of the OCG (and hence departments) did not coincide with those of either the Treasury Board Secretariat—the central finance agency responsible for resource coordination—or the cabinet (Campbell 1983, pp221-222, 227; Savoie 1990, pp114-116; Segsworth 1993, pp100, 108). As a consequence, institutional jostling and ministerial indifference ensured that evaluation was never integrated into the decision making processes for the expenditure budget.

The second level of the literature indicates that efforts to strengthen the capacity of finance agencies to exert influence in the budget process have been hampered by the weakness of the link between evaluation policy and budget review procedures. The question is, what explains this infirmity? Campbell suggests that the two processes are "inherently incongruous"—there are "clear limits" on the capacity of institutional arrangements for improving coordination between the two. At its base the ability of management techniques to influence expenditure review is constrained by the fact that political considerations will "normally outweigh data from effectiveness evaluations" (Campbell 1983, pp204, 347-348). He argues that the location of the resource allocation function does not influence the potency of expenditure review as much as the impact of policies designed to improve management (for our purposes, program evaluation policy). For Campbell, the design problem is that management policies are influenced by the relationship between resource allocation functions and economic policy setting functions, since the former are often neglected in departments predominantly concerned with economic and fiscal policies:

> [s]plitting off detailed expenditure review and brigading it with units responsible for management policy would give a context to the latter that it simply lacks when not directly connected to some specific part of expenditure review (Campbell 1983, pp345, 349-350)

This helps explain the problems with budgetary reform in both the United Kingdom and Canada, since in the former developments in program evaluation policy were often given short shrift in relation to the setting of fiscal parameters of macroeconomic objectives, while in the latter the linkage between evaluation policy and expenditure review was severed by housing evaluation policy in an agency separate from both. As a consequence, because of the inevitable inter-agency rivalries the information from analytical techniques was not incorporated into expenditure review (Campbell 1983, pp221-222, 227). We should note, however, that Campbell's distinction is somewhat artificial—ultimately the location of evaluation policy is important precisely because in order for it to influence expenditure review it must be seen by spending departments to be a component of "the one solvent of government—having to pay the bills" (Heclo & Wildavsky 1974, p278).

From this brief survey of the comparative literature a number of fundamental questions present themselves for this thesis.

First, to what extent does the introduction of program evaluation within the Australian system represent an enduring change in the rules of the game of "policy management"? Efforts to integrate policy making and expenditure management in other governments have, almost without exception, failed in their primary task of recasting the complexion of central budgetary and policy coordination processes. Instead the new rules and routines have most often either been absorbed into the existing—dominant—rules of the game, or have made the rules more complicated and unworkable (cf Self 1980, p36). In other words, we are interested in gauging the extent to which evaluation has transformed the rules in line with changes in the playing conditions—that is, the scope of policy coordination questions (Painter 1981, pp279-280; Painter 1987, pp192-193).

Second, given the changes in the complexion of coordination questions facing central agencies—and budget agencies in particular, which are now much more likely to be oriented towards the substance of public policy—how has the role of these central agencies evolved following the introduction of program evaluation? What are the limits to these roles, and what is the potential for “policy overstretch”? Furthermore, what are the consequences of policy overstretch? Can they be avoided? How?

In addressing both sets of questions, this thesis will focus mainly but not exclusively on the Commonwealth Department of Finance. It is directed primarily at examining and assessing the influence of one policy analytical technique—program evaluation—on the work of central agencies and their policy management functions. As the next two chapters will describe, program evaluation policy in the Australian federal government, while generally following broad international trends surveyed in the previous chapter, has taken on a number of distinguishing characteristics. As we will see, although the Labor Government’s evaluation strategy was catalysed by the economic crisis of the mid 1980s, the policy itself had a long gestation period and this must be taken into account when explaining in detail the development of the current policy.

CONCLUSION

This chapter has sought to construct a framework for analysing the role of program evaluation in Australian central government. It examined the nature of the central coordination problem in government and argued that changes to the policy environment had reoriented the scope of these questions from procedure to the substance of policy. It was also argued that the role of central agencies could be seen in terms of a policy management function that was principally concerned with three aspects of the newer type of coordination problems. The role that program evaluation plays in the new policy management environment presented itself as the open question. Has evaluation recast the way that the policy management function is approached? Does evaluation lead to the over-extension of central agency policy and resource coordination capabilities? These are the issues that are addressed in the chapters that follow.

Chapter Four

Evaluation In The Policy Stream 1976 - 1986

The purpose of this chapter and the one following is to trace the development of program evaluation within Australian government. The antecedents of the current evaluation strategy have been briefly sketched elsewhere (Corbett 1991, pp3-9; Ryan 1992, pp65-66), but a detailed analysis of the long administrative design process informing the strategy deserves more considered treatment. This process will be considered in two stages. The present chapter will examine the early period of development, taking the accountable management framework outlined in 1976 by the Royal Commission on Australian Government Administration (RCAGA) as its starting point. It will demonstrate that the major obstacle confronting the development of an evaluation policy has been establishing the primary purpose that it should serve. Although there is a stream of public policy running through the evaluation strategy that can be traced back to the accountable management framework, there are also a number of complicating twists and turns in the flow. Accountable management marked out changed relations between the institutions of Australian government, and program evaluation, with its new undefined review criterion of effectiveness, was quickly recognised as a crucial political resource within this policy arena. As a consequence, distinct and often competing purposes for program evaluation emerged in this period that reflected intense institutional jostling by political executives, parliament and the bureaucracy.

A Frame Of Reference: The RCAGA And Accountable Management

A central argument of this chapter is that the current evaluation strategy can only be fully understood in terms of the frame of reference which informed its development. This can be found in the accountable management framework constructed by the 1976 *Report* of the RCAGA. It is readily conceded that the

Report made somewhat equivocal pronouncements, but any ambiguity can be readily attributed to the width of its terms of reference which sought to incorporate the concerns of diverse public service interests and ensured that the "inquiry was given no precise agenda for reform" (Hawker, Smith & Weller 1979, pp242-243; cf Spann 1977, p81). In any event, the Commission was able to produce an agenda of which two aspects will be pursued here. The first is the Commission's framework for accountable management and improved central coordination. The second consists of the more specific structural design questions relating to an evaluation function.

For a report that was greeted with a curious mix of indifference and outright hostility (Thompson 1979, p80; Nethercote 1984, p17), the RCAGA has maintained an enviable reputation as the touchstone of Australian government administration. This attests to the success of the Commission as "interpretative authority", an authority which has been achieved through its shaping of the Australian administrative reform agenda and its "long term impacts on institutions and policy outputs" (Stone 1994, pp256-8).

To begin, the RCAGA must be placed in context or, more precisely, firmly located at an important juncture in the development of contemporary Australian government. Arguably, this juncture marks the end of an absolute faith in—and the beginning of a sustained questioning of—the inherited principle of individual ministerial responsibility which so comprehensively defined the terrain of public accountability. The view was prevalent that, as an organising concept, ministerial responsibility had outlived its usefulness because it no longer reflected political and bureaucratic reality, while the "effectiveness of cabinet has suffered from the 'structural disintegration of the decision making process'" (Emy 1976a, p24; 1976b, p54). As a consequence, the RCAGA's administrative reform project was to reflect a distinct focus on enhancing political control of departmental activities:

> The fundamental task is to integrate the authority which comes from popular election with that which derives from professional knowledge and experience, while upholding the principle of ultimate political control. If the two sources of authority are to be so integrated, the roles of minister and official must be seen as complementary (RCAGA 1976, para 3.4.3, p43).

Political control was framed predominantly as ministerial in nature, although the "principle" sought also to acknowledge wider responsibilities to parliament and the public. To this end, the RCAGA was directed at two purposes which were designed to be mutually re-inforcing. In the first place, the priority setting and co-ordinating capacities of the political executive—cabinet—with respect to government departments had to be strengthened. In essence this meant increasing the clarity of government objectives and the specificity of ministerial decisions "so that ministers will in fact be conscious of what they are deciding and significant decisions of a political character will not, by default, be made by officials" (RCAGA 1976, para 3.3.5, p38). The instrument for this was a refined forward estimates process. Second, the guiding principle of ministerial responsibility was open to reconstruction by a redefinition of the role of the public service. Ministerial responsibility limited the role of the public service to the criteria of the policy-administration divide which, many were arguing, no longer held either descriptive or prescriptive value. In effect the public service defined itself only by reference to the tenets of ministerial responsibility—neutrality, anonymity and instrumentality—when what was required was an explicit conceptual recognition of the public service as "exercising some powers in its own right" (RCAGA 1976, para 2.1.9, p13; or, alternatively, having a "separate constitutional persona", see Emy 1976a, p63; Spann 1977, p83).

The Commission's pragmatic approach was to supplement ministerial responsibility with a new system of administrative accountability—accountable management. Accountable management attempts to acknowledge the responsibility which public servants assume in policy making and limit this through a rudimentary results accountability system:

> The theory of the Westminster system asserts that the minister is wholly responsible for all actions in his department, but in fact much responsibility lies with officials. It is important that this be acknowledged, the nature and extent of the responsibility be clarified as far as possible, and procedures established to assess performance and to provide that those responsible at all levels will be accountable for their performances (RCAGA 1976, para 3.4.1, p42).

Hence, the two arms of reform as contained in the RCAGA's construction of accountable management were mutually reinforcing: the policy role of the public service is recognised and limited by internal accountability for performance but, at the same time, the position of the political executive is not diminished because the central co-ordinating and priority setting capacity of cabinet is bolstered; that is, there is increased participation by both spheres in the policy and administrative process while the increased policy role of the administration is to be contained within set parameters (cf RCAGA 1976, para 3.3.5, p38).

Intuitively, this version of accountable management would appear to generate considerable potential for tension between strong centralising and decentralising trends. This last construction, however, misinterprets the Commission's prognosis of what ailed existing principles of political control—in essence, both failing political control and administrative inadequacies could be attributed to *coordination incapacities* at the centre. As a consequence, the prescriptions contained in accountable management must be seen as radiating from the centre for the primary purpose of strengthening the centre. Given this, it is important to sketch its mechanics in some detail.

As a response to perceived deficiencies in the principle of individual ministerial responsibility, accountable management is essentially a form of institutional renovation—it attempts to redefine political accountability by integrating political and bureaucratic authority while "upholding the principle of ultimate political control", a formidable task. For the Commission accountable management revolved around securing the prerequisites of efficiency. These prerequisites were, in turn, predicated on increasing "coordination" at the centre. The process of coordination is concerned with the question of how to make effective the discipline of constraints facing the government (such as the interdependencies of the collective responsibility of cabinet and the limits imposed by external economic, social and political considerations) (RCAGA 1976, paras 11.1.1-5, p355). Importantly, coordination is only effective when it is based on "voluntary acceptance of discipline" rather than the externally (Treasury) imposed variety which characterised the existing process (RCAGA 1976, para 11.1.10, p357). To satisfy these conditions, coordination involves the facilitation of more active participation by *both* ministers and their departments in the determination of the nature of constraints and the priority of government objectives. The primary instrument for this task is the forward estimates budget process.

The forward estimates process was designed to encourage voluntary acceptance of discipline by producing "generally acceptable" assessments of the resources and constraints facing government and providing for the participation of ministers and their departments in the negotiation process. The forward estimates themselves would be policy guidelines formulated by specially convened cabinet committees serviced by working groups of senior officials drawn from central agencies (RCAGA 1976, para 11.2.1-14, pp357-62).

These policy guidelines would ensure that

> the establishment of objectives, their embodiment in programs and the determination of priorities between them are *essential functions which link the political and administrative aspects of government* (RCAGA 1976, para 3.3.1, p36, emphasis added).

Arguably, this link represents the core of accountable management. As discussed earlier, the Commission sought to redefine political accountability by recasting the role of the public service. In examining alternative locations for responsibility the Commission noted that one

> approach, which comes closer to practice in the private sector, is to acknowledge the managerial responsibility of officials, but also to prescribe the means by which their performance will be assessed and reported upon in ways which will influence ministerial judgment about their professional standing and future (RCAGA 1976, para 3.4.6, p43).

As a consequence this approach sought to create an incentives structure for accountable management explicitly linking career advancement and increased managerial responsibility. Hence, the purpose of the forward estimates was to provide a framework for this responsibility by translating policy objectives into programs to be “performed” by departments. Efficiency would be increased by promoting the “entrepreneurial function” and devolving decision making authority to lower levels of management, while at the same time ensuring that management would be held accountable for their decisions through the regular monitoring of departmental performance (RCAGA 1976, para 3.2.3-11, pp34-6). For a number of reasons, as the next section will show, the envisaged mechanism of review was the efficiency audit.

Of course, the application of accountable management to government administration had its antecedents, but unlike the RCAGA which innovatively attempted to tie accountable management to a broader consideration of the public service and political control, these earlier tracts were somewhat more partial in their consideration of the concept (Spann 1977, p86). An early formulation—and one which certainly resonates with the tenor of the RCAGA—can be found in the Canadian Royal Commission on Government Organisation (the Glassco Commission) of 1960-62. Also initiated by a party long denied government and suspicious of its public service, the Commission developed what it termed a "concept of management". This pictured effective government management as falling into two categories: the administration of departmental operations and the central direction and coordination of government activities as a whole (Canada 1962, p48). Wholly unimpressed with existing arrangements, Glassco was concerned with delineating these two tasks by simultaneously strengthening central capacity to integrate budgetary and policy decision making and reducing central incursion into departmental operations. The Commission's prescription was to overhaul the budget process by introducing program budgeting, a system of forward estimates and the creation of a task-specific coordinating unit to be located within the Privy Council Office (Canada 1962, pp97-101). The intention was for this unit—which eventually emerged as the Treasury Board Secreatariat—to assume a leadership role and guide managers in departments. At the same time, however, the Commission argued that the Board should "let the managers manage" by delegating the "power of decision" and making them accountable for the management of resources (Canada 1962, pp33, 103).

Here then is accountable management in substance if not in form, drawn directly, as the Commission notes, from "management concepts and techniques of internal control developed in the private sector of the economy" (Canada 1962, p51). This last point indicates that the Glassco formulation of accountable management was shaped by the remedies offered by private business practice and facilitated by a "strong private sector presence" on the Commission (Savoie 1994, p61). It also highlights the primary shortcoming of this rudimentary framework, since according to critics it "failed to address the operation of Cabinet and Parliament" by neglecting the question of *how* management would be held accountable both within government and to parliament (Savoie 1990, pp127-8; Savoie 1994, p75).

Likewise, the British Fulton Committee of 1966-68 conceived accountable management as a purely managerial prescription, following practice in what it called "progressive industry" (Cmnd. 3638 1968, para 153, p52). It was meticulously defined as

> holding individuals and units responsible for performance measured as objectively as possible. Its achievement depends upon identifying or establishing accountable units within government departments—units where output can be measured against costs or other criteria, and where individuals can be held personally responsible for their performance (Cmnd. 3638 1968, para 150, p51).

This orientation was not surprising. Again, the impetus for this recommendation came from analysis undertaken by a management consultancy group. The recommendation itself was a “pragmatic one”, having the sole intent of providing support for the Committee’s primary agenda of remedying the stultifying effects of the civil service class system in the United Kingdom (Fry 1993, pp71, 133). References to ‘personal responsibility’ indicate that it was fundamentally about improving efficiency through better career structures. As a key adviser on the consultancy group was later to remark:

> what we wanted was a fairly benevolent system of accountable management . . . with minor sanctions but a fair bit of preferment (Fry 1993, p68).

In the end, the Fulton Committee’s consideration of accountable management constituted only a tiny fraction of the recommendations it made. The most important proposals were directed at “the type of civil servant needed to undertake the tasks of modern government irrespective of the precise way Whitehall might organise itself to carry out these tasks” (Fry 1993, p28); that is, it was, first and foremost, occupied with diluting the ‘cult of the amateur’. In no small part this can be attributed to the driving force of a lead group of committee members who perceived the committee’s report in terms of a “truly radical document” designed to promote quite specific cultural change, and generated undoubtedly by a remit which explicitly excluded examination of the doctrine of ministerial responsibility and the machinery of government (and, by extension, “the power of the Civil Service”) (Fry 1993, pp22, 28-9; cf pp10-11). This last obstacle, of course, was one with which the RCAGA did not have to contend.

Clearly then, the RCAGA project can be distinguished from its northern hemisphere antecedents on two important counts. First, the RCAGA was born with terms of reference which provided “an almost unrestricted field for investigation” (Hawker, Smith & Weller 1979, p249). Ultimately, this allowed the RCAGA to pitch its discussion at a higher level—its conceptual model for accountable management attempted to supplement aspects of the doctrine of ministerial responsibility by tentatively acknowledging reality (that is, the policy responsibilities of the public service). In stark contrast, both Glassco and Fulton were restricted by relatively narrow terms of reference which produced somewhat constrained considerations of accountable management as essentially an organisational prescription. As a consequence, both accountability and the role of performance evaluation received comparatively short shrift. Second, membership of the RCAGA did not include representatives from private business (Hawker, Smith & Weller 1979, p247). At this most visible level, at least, a reasonable extrapolation might be that, unlike its predecessors, the RCAGA version of accountable management was less likely to reflect an uncritical absorption of business management techniques—it was “a broad concept of control rather than simply a businessman’s approach to government” (Emy 1976a, p49). Indeed, accountable management was founded on a redefinition of public service responsibilities that was tinted with what Self has termed the “social agenda” of reform, or an attendant concern with building in increased responsiveness to community needs (Self 1978, pp313-314).

THE RCAGA TREATMENT OF EFFECTIVENESS REVIEWS

As the previous section outlined, the RCAGA version of accountable management stressed that the devolution of managerial decision making authority must be accompanied by commensurable responsibility and accountability for performance. Working from this proposition the Commission proposed that the primary mechanism of accountability would be a regular program of efficiency reviews to assess departmental performance (RCAGA 1976, para 3.6.1, p46). Given the Commission’s frame of reference—the enhancement of political authority—the crucial question quickly became where to locate this function. It canvassed several possibilities. The Public Service Board (PSB) already had a number of responsibilities in relation to promoting economy and efficiency in the public service deriving from section 17 of the 1922 Public Service Act. However, the Commission noted that the PSB had shown “great reluctance” in assessing the performance of departments and that this was likely to continue given the management advisory role played by that agency. Other options included the creation of a new agency—the Office of Policy Analysis and Administrative Management (OPAAM)—in which, however, the Commission saw “little merit”, or extending the audit function of the Auditor-General to embrace both compliance and efficiency. The last alternative presented itself as the “most appropriate” (RCAGA 1976, para 3.6.17, p49; see also recommendations in PSB 1975, pp13-15).

Of necessity, efficiency reviews excluded considerations of effectiveness. For the Commission, efficiency reviews could be distinguished both conceptually and operationally from what it called “program effectiveness reviews”.

Assessing the effectiveness of government programs “would require essentially political judgments in which, in the Commission’s view, the Auditor-General should not be involved” (RCAGA 1976, para 3.6.18, p49). Instead, effectiveness reviews would be “most appropriately” arranged within the Department of the Prime Minister and Cabinet (DPMC) where such reviews “should be linked with the development of new programs and government priorities” (RCAGA 1976, para 3.6.19, p49).

As we have already seen, a principal concern of the Commission was to strengthen the priority setting and coordinating capacities of the cabinet through the forward estimates budget process. It was envisaged that DPMC would play a pivotal role in ensuring that the forward estimates guidelines reflected the content of government policy. To facilitate this role, the *Report* recommended that a centrally placed “policy unit” be created within PMC charged with a responsibility

> to evaluate the effectiveness of ongoing programs against the general philosophy, policies and priorities of the government (RCAGA 1976, para 11.5.21, p385).

This type of evaluation would constitute an “intensive analysis of selected programs chosen because of their political or financial importance” and directed purposefully at augmenting the policy awareness of ministers. It was, accordingly, contemplated as an “essential” role in the preparation of those guidelines for the forward estimates process which identified the major “political issues” associated with resource allocation (RCAGA 1976, paras 11.5.18-19, pp384-85).

The placement of responsibility for effectiveness reviews within DPMC’s sphere signalled clearly the Commission’s intentions for program evaluation as an important component of the information flow required for improved central coordination (that is, in determining the objectives and priorities of the government and assisting with the clarification of the responsibilities of ministers and departments within a more collegial construction of collective responsibility). Since coordination was equated with the budgetary process (RCAGA 1976, para 3.3.3, p37) program effectiveness reviews—to be undertaken on an ad hoc basis—were to be directed towards determining policy priorities in resource allocation. In effect, and especially for the purposes of accountability, the RCAGA’s model of accountable management had made a crucial distinction between efficiency and effectiveness—the accountability cause was to be served by efficiency audits, which the Commission presumed to be less problematic, while effectiveness reviews were explicitly quarantined from questions of public accountability. They were to serve higher (budgetary and policy strategic) purposes.

The RCAGA vision for effectiveness reviews was not without its detractors. One prominent contributor to the Commission's extensive research effort prepared a timely and well crafted critique (Cutt 1977). According to this account, for accountability purposes efficiency could be defined in a hierarchy of increasing "sophistication and potential usefulness": efficiency I (financial audit) demonstrates the propriety and legitimacy of expenditures; efficiency II (efficiency audit) measures the ratio of program outputs to program costs; and efficiency III (effectiveness audit or program evaluation) assesses the degree of success in the achievement of program objectives. As Cutt notes, the RCAGA readily adopted efficiency III:

> For the purposes of this Report, effectiveness is one of two distinguishable elements in efficiency. Effectiveness is concerned with the relationship between purpose and result. Thus an action or program is effective if it achieves the purpose for which it was initiated. But efficiency involves additionally a consideration of the resources used in achieving the result. A program is efficient only if its effectiveness is achieved with an economic use of resources. A program is efficient only if its effectiveness is achieved with an economic use of resources. Efficiency is therefore also concerned with the relationship between resources used and the results achieved; between 'input' and 'output'. *It comprehends both economy in this sense and effectiveness* (RCAGA 1976, para 3.1.4, pp31-2, emphasis added).

Although it also proposed the extension of the Auditor-General's audit function into efficiency II, Cutt found it "difficult to resist the conclusion" that the Commission had "failed to make a serious commitment to the concept of a comprehensive (effectiveness) audit implied in its own definition of efficiency" (Cutt 1977, p343; cf Dillon 1985, pp254-5). Further evidence of such vacillation was to be found in the proposal to locate the responsibility for efficiency III reviews in DPMC. For Cutt it was imperative that a specific agency—similar to OPAAM—be responsible for program evaluation "on a continuing, systematic basis" and that a policy unit within DPMC "cannot contribute seriously to, and be a functioning component of, a comprehensive approach to analysis" (Cutt 1977, p346).

If this first line of criticism took the Commission to task for timidity, a second was to reprove it for its apparent naivety. In order to redefine notions of public accountability in terms of increased political control, the RCAGA reasonably started from the premise that central policy coordination capacities should be strengthened. For some commentators problems emerged when the RCAGA decided to equate coordination with the development of forward estimates and thereby “pinned its hopes on the budget process” (Painter & Carey 1979, p95). This was bemoaned as misplaced faith, if only because of the political fact that

> the budget process does not encompass all aspects of policy coordination, in particular those aspects that arise from the need to respond to unforseen demands and to adjust to changing policy circumstances (Painter & Carey 1979, p96).

At the same time, its attempts to redefine political control in such a way as to grant the public service more explicit policy responsibilities, meant that the Commission’s accountable management prescription was likely to extend rather than alleviate coordination problems because it conveniently sidestepped the unruliness of bureaucratic politics (Smith 1978, pp185, 190; Painter & Carey 1979, p93).

In sum, the RCAGA treatment of effectiveness reviews was somewhat problematic. On the one hand it is unsatisfactory because it fudged its consideration of the efficiency-effectiveness separation. From its own definition above, efficiency includes both economy and effectiveness, with the latter introducing all the difficulties associated with judging the achievement of objectives and the appropriateness or fit “between purpose and result”. Although accountable management was constructed by the RCAGA in such a way as to articulate a new responsibility of public servants for effectiveness, the RCAGA was not prepared to make them publicly accountable for this since the accountability scheme was so designed—principally, by locating the review function with the Auditor-General—as to explicitly excise consideration of such ‘political’ questions (although even here the Commission harboured reservations about a strict separation between efficiency and effectiveness noting, in relation to the proposed efficiency audits, the possibility of “overlap” since the “results or benefits contemplated will generally be a complex and sometimes competitive set of objectives” (RCAGA 1976, para 11.4.4, p376; para 11.5.20, p385)). Therefore one could be forgiven for concluding either that the RCAGA’s version of accountable management had, by locking itself into the conventional politics-administration divide, failed in its task of redefining public accountability or, given its obfuscating definition of efficiency, somehow sought to surreptitiously hand the Auditor-General a remit to examine questions of effectiveness. At the same time, however, there is a compelling logic to the Commission's reasoning for confining program effectiveness reviews to informing the central task of setting government-wide priorities. Since this was essentially a budgetary process, questions of effectiveness—and program effectiveness reviews—were best tackled at this level.

Clearly then, at the time of the *Report* of the RCAGA a number of tensions were pulling evaluation policy in different directions. While the RCAGA viewed evaluation through the lens of accountable management, it did not envisage effectiveness reviews as a tool of accountability. Rather, its primary task was to assist in the determination of priorities for resource allocation and, because of its policy sensitivities, evaluation was best placed in a central co-ordinating department. Outside observers were less sanguine about the efficacy of this organisational location and pushed for a separate agency to systematically evaluate government programs. Transcending all of this, however, was a confusion over the distinction between the efficiency and effectiveness of government programs.

Finally, the *Report's* systemic proposals for accountable management were also not universally well received. Even though its proponents claimed it was "a direct response to the problems of ministerial responsibility" (Emy 1976a, p47) accountable management was assessed by several commentators to be a patently inadequate response. The *Report* runs into trouble by way of conflict between the competing models it develops. It critiques the politics-administration dichotomy by suggesting its dilution—via increasing joint involvement of ministers and officials in policy and implementation work—but this grates with the thrust of accountable management, which attempts to acknowledge the responsibility of *officials* and hold them accountable for performance. At the same time the *Report's* distinction between questions of efficiency and effectiveness paradoxically reasserted the politics-administration dichotomy. The final product was judged to weaken ministerial responsibility (Spann 1977, pp82-4). Of course the question of developing a notion of political responsibility which incorporated public servants with an independent constitutional status remained open. While the RCAGA served to point a direction for answering this question, program evaluation was only a small distinguishing mark on the map.

A final criticism was that accountable management failed to resolve the attendant questions of to *whom* public servants were to be held accountable and how. This issue was brought into relief by the *Report's* sole reliance on efficiency audits as the accountability mechanism for public service responsibilities for program performance. On the face of it, this proposed that public servants were ultimately accountable to the Auditor-General and, by extension, the parliament. The problem with this rendition of accountability was that it was likely to categorically preclude most, if not all, parliamentary consideration of program performance because the Auditor-General's efficiency audit mandate explicitly denied access to questions of 'program effectiveness' (RCAGA 1976, para 3.6.18, p49; Reid 1976, pp326-7). As a consequence it was inevitable that accountable management, which envisaged public accountability predominantly in terms of performance and results, was to locate responsibility for effectiveness reviews—program evaluation—within the executive. Indeed for one commentator the *Report*, far from representing a fundamental reconsideration of existing structures, merely constituted

> a rationalization of existing trends in Australian government. It is based on *post hoc, propter hoc* reasoning—that the existing state of government in Australia (as in Britain), the decline of Parliament, and of parliamentarians in the governmental process—justifies organising alternative methods of accountability and responsibility; methods that will not arrest the decline in parliamentary scrutiny and control but will accelerate it (Reid 1976, p327).

This was one of the core tensions which plagued—and continues to trouble—the *Report's* accountable management framework. It was also this scenario which several parliamentary committees sought to address towards the end of the 1970s.

In the aftermath of the *Report*, other commentators correctly observed that many of the Commission's proposals (including those relating to program effectiveness reviews) did not represent a "radical break with the past" since they were "consistent with ideas recently emanating from the central agencies" (Nethercote 1977, p101). The PSB provides a case in point. The Board was acutely aware of its role in the promotion of efficiency throughout the public service, concentrating on management improvement at the central level through, for instance, the program of Joint Management Reviews (JMRs) which commenced in 1974-75 (PSB 1977, p96; 1978, p31). However, in parallel with the RCAGA the Board also noted how

> there is increasing recognition in public services throughout the world of the need to develop as far as possible techniques for measuring productivity and for assessing outputs and the effectiveness of programs implemented in response to community needs (PSB 1976, p5).

In particular it stressed the urgency of measures to “strengthen arrangements for review of overall performance of government programs” (PSB 1976, p17; 1975, pp4-5).

In November 1976 the Fraser Coalition Government announced new administrative arrangements which were to have repercussions for the evolution of program evaluation. The economic policy analysis and advice function of the Department of the Treasury was separated from its financial management and control activities, with the latter being transferred to the newly created Department of Finance (hereafter Finance) in December of the same year. According to Prime Minister Fraser the changes were designed to

> provide for more effective management of the business of Government, and to strengthen the Government’s decision making processes. Particular emphasis is being laid on improving capacity to service the Government’s requirements for forward planning, priority setting and the strategic planning of Government initiatives (Commonwealth Parliamentary Debates (CPD), House of Representatives, 18 November 1976, p2898).

These changes, he explained, were made in light of “careful consideration” of the recommendations contained in the *Report* of the RCAGA, although they are just as likely to have reflected other, more immediate political considerations, such as breaking Treasury’s monopoly on economic policy advice (Weller 1989, pp73-74; Warhurst 1988, p331). Nonetheless, associated changes did herald a new rhetorical commitment to program evaluation. There would be a “strengthening” of resources available to DPMC for evaluating the effectiveness of programs on a regular basis and “where appropriate these evaluations will be followed by detailed reviews and the development of modified or new programs” (CPD, House of Representatives, 18 November 1976, p2899). These reviews were to be “closely associated” with the accelerated development of the forward estimates system.

The creation of Finance was to have long term consequences for the shape of program evaluation policy via its gradual reworking of procedures for financial control. (This development receives more sustained treatment in Chapter Six.) A major implication of the split was that Finance's control of public expenditure was more insulated from the macroeconomic considerations which were now solely the preserve of Treasury. Inevitably, Finance inherited the intensely hierarchical structure which characterised Treasury operations, and which manifested itself through variable modes of control employed by supply divisions, as well as the so-called "accountant's mentality" that line departments found so abrasive (Weller & Cutt 1976, p39). The power of the Treasury as the pre-eminent coordinating department could be directly attributed to its mutually reinforcing responsibilities for control of public expenditure and macroeconomic management. The old Treasury was marked by the subordination of supply division activities to the General Financial and Economic Policy (GFEP) division which undertook the department's primary responsibility for advising on macroeconomic policy management (Weller & Cutt 1976, pp29-31, 36-37; Weller 1977a, pp31-2).

Hence, the principal issue facing the new Department of Finance was the extent to which it could overcome this kind of structural and procedural legacy. Clearly, there existed considerable potential for change in expenditure control and policy coordination—for instance, by ameliorating the "depressing" effect that GFEP had on supply divisions so as to develop more systematic techniques for the evaluation of program expenditures. Indeed, it was argued that only "when financial control is an end in itself, rather than a means for another section of the department" could such change occur (Weller 1977a, p33). Given that any move to effectiveness reviews required an enhanced technical capacity, it was likely this would first emerge within an agency more attuned to expenditure management issues, as Finance was. As a consequence, overcoming the procedural inertia affecting its expenditure control functions called for a change of "style" which, in turn, required Finance to define its own functional identity (Weller 1977a, pp35, 39). Re-emphasising the linkage between program evaluation and resource allocation would eventually present itself as one route to this objective.

The thrust of these changes carried forward the RCAGA evaluation banner. The stress was on planning and priority setting, and feeding evaluation results into this resource allocation process. There were developments. By the end of the 1970s Finance was able to report its participation in several *ad hoc* program effectiveness reviews (DoF 1979, p4). Likewise, from early 1976, DPMC boasted a projects division that had amalgamated a number of the policy development units established by the coordination conscious Whitlam Labor Government (Walter 1992, pp35-38). It consistently listed among its responsibilities "the coordination of program review activity" and the "evaluation and development of programs and policies in selected major areas of government interest" (DPMC 1979, pp5, 17; Commonwealth of Australia 1977, p218; 1978 p223). However, the arrangements for effectiveness reviews were "still at an early stage of development" and although the initial practice was to conduct reviews in the period in the *run-up* to the budget, there is no indication of the volume, or use made, of this review activity. With the appointment of Geoffrey Yeend to the secretaryship of DPMC in early 1978, the new emphasis on review activity was played down in deference to a more conventional view of what was administratively feasible (Hawker, Smith & Weller 1979, pp121-122; Walter 1992, pp38-39).

PARLIAMENT AND EVALUATION

Without doubt the RCAGA had led to some ferment of ideas on the development of program evaluation within Australian government. However, progress was slow. It is not surprising that the impetus for further reform would come from outside the administration. Against a background of growing public malaise about its role in the processes of governance, parliament embraced the idea of program evaluation as one way of strengthening and expanding its role of scrutinising the executive. This parliamentary interest not only reflected RCAGA concerns to improve the volume and quality of evaluation activity but, in a significant development, sought also to extend responsibility for this activity from the confines of the executive to the open spaces of parliament. Working in separate houses, but with an awareness of the other's line of inquiry, two parliamentary committees were to add another layer of expectations for program evaluation.

In the late 1970s a number of Senate Estimates Committees were becoming increasingly disaffected with the way in which expenditure on government programs was being justified (Uhr 1990a, pp103-5). In particular, those committees principally concerned with health and social welfare programs were critical of the way in which programs were "consistently" funded in the absence of explicit objectives and subsequent assessment of effectiveness (Senate Estimates Committee D, 28 April 1977, p61; Senate Estimates Committee C, 2 May 1978, p49). The little evaluation which was undertaken was *ad hoc* and thus usually failed to make an integrated contribution to either the development of new policy or reconsideration of existing policy. This laxity not only affected the capacity of government programs cost-effectively to deliver the health and welfare services required by the community, it also eroded the ability of Estimates Committees adequately to scrutinise the appropriateness of government expenditure. As a consequence, it was resolved by a number of committees actively to encourage a system of program evaluation within the administration so as to render the estimates process “truly effective” (Uhr 1990a, p104).

By way of contrast, the House of Representatives Standing Committee on Expenditure (HRSCE) provides early evidence of a determined, yet short-lived excursion by *parliament* into the evaluation of policy effectiveness (Walsh 1991a, p21). The HRSCE was established in April 1976 to “inquire into and report on any question in connection with public expenditure which is referred to it by the House” (Resolution of Appointment, CPD, House of Representatives, 8 April 1976, p1496). Expectations for the success of the committee were high. In the first place it was

> [p]otentially . . . a very powerful committee because, as an ongoing concern it can investigate the activities and efficiency of departments, challenge their competence and publicly castigate them for their failures (Weller 1976, p37).

In the second place it enjoyed—at least initially—the support of a prime minister who had proposed its creation as a spur to improved parliamentary oversight (CPD, House of Representatives, 8 April 1976, pp1497, 1499; Reid & Forrest 1989, pp163, 376-377). It was inevitable that, to some degree, the functions of the HRSCE overlapped with those of the Estimates Committees and, perhaps for this reason, the Committee sought to carve out a distinct role for itself. In reviewing its first year of operation, the Committee was able to distinguish three primary functions: the examination of economy and efficiency in the use of public funds, the evaluation of program effectiveness and results, and the examination of processes for the management and control of public expenditure (HRSCE 1977b, p8). While the inaugural Chairman of the Committee, Ransley Garland, succinctly expressed the commonality of these roles as “value-for-money considerations” (CPD, House of Representatives, 24 August 1976, p462), the functions were also concerned with different and contentious aspects of the same considerations.

This was duly reflected in the first two reports of the Committee where the adequacy of government policy was routinely scrutinised (HRSCE 1977a; 1978). Although the Committee accepted its operational guidelines—that it "take stated Government policies as given" and confine investigation to "implementation"—it was also of the view that "policy matters" were more than likely to be touched on during the course of a thorough examination (HRSCE 1977a, px; Operational Guidelines, CPD, House of Representatives, 8 April 1976, p1499). To cite an example, in relation to a disastrously unsuccessful forestry program, the Committee was inclined to criticise "unrealistically ambitious" objectives and dispute the very rationale for the program (HRSCE 1978, p6).

This nascent policy scrutiny function for parliament was more fully articulated in the HRSCE's report *Parliament and Public Expenditure,* which has been aptly depicted by one commentator as a "forgotten classic of Parliamentary confessional literature" (Uhr 1990a, p86). That the Committee was disclosing its resignation to widely held fears for parliament was made abundantly clear by its second chairman when tabling the report:

> This report is a product of its time [and] simply had to be written today because of the growing concern that the elected Parliament is a weak and weakening institution. It is a report that had to be written, now that talk of parliamentary reform is in the air, so as to provide a sharper focus for some of the debate (Mr Kevin Cairns, CPD, House of Representatives, 3 April 1979, p1406; cf Reid & Forrest 1989, p376).

But, as this suggests, the report was also a manifesto for change, and program evaluation was presented as the key to that change.

For the HRSCE the major issue facing parliament was how to enhance its influence on the public expenditure process. The problem was characterised as the absence of both adequate procedures to affect executive budget formulation and of appropriate information for review purposes (HRSCE 1979, piv). On the first problem, the Committee reaffirmed the course taken by the RCAGA towards forward estimates and a more transparent budget process. The information needs of parliament for review was a more delicate issue. In the first place there was an acceptance that most review work was undertaken within departments, classified as "internal working documents" and therefore of negligible accountability value (HRSCE 1979, para 57, p19). It was also acknowledged that the spectrum of review ranged from compliance through to efficiency, effectiveness and "policy and priority", but as review work moves along the spectrum towards the policy end, it "becomes more important, relatively speaking, and more political" (HRSCE 1979, para 52, p18). And, although the Committee conceded that it was the only parliamentary committee that sought to evaluate effectiveness "on a continuing basis", herein lies the rub. For while it might be desirable for parliament to have a policy evaluation function of its own, more concrete moves in this direction were likely to dissolve the bi-partisan compact upon which contemporary parliamentary—that is, committee system—review rested (HRSCE 1979, para 59, p19; see also Uhr 1982, pp42-3).

The Committee's solution was to sharpen a distinction, perhaps artificial, between the evaluation of program effectiveness and the evaluation of "strategic priorities decided by Cabinet" (HRSCE 1979, para59, p19; Walsh 1991a, p22). The effectiveness of programs—that is, the extent to which program outcomes were achieving stated objectives—was acceptable scrutiny, where as questioning the relevance of government policy and priorities, as had been attempted in some of the earlier reports by the Committee, was now strictly out of bounds. This more limited program evaluation function of parliament did, however, call for "specific information" in the form of program statements which linked cost information with "effectiveness measures" (HRSCE 1979, para 60, p19).

In many ways *Parliament and Public Expenditure* represented a last ditch attempt by those parliamentarians concerned with the adequacy of processes for the control of public expenditure to see the most crucial recommendations of the RCAGA accepted by government. It is also instructive for the reason that it saw parliament as an evaluator of program effectiveness, an arrangement which has largely fallen to the wayside. For the Committee, program evaluation was *central* to enhanced scrutiny of the public expenditure process and it, in turn, could only be effective with the development of program budgeting. Indeed,

> [t]he ultimate purpose of program statements is to enable programs to be evaluated. They would provide an information base for the Parliament to discuss public expenditure and priorities in a more meaningful way, and permit systematic scrutiny of programs by committees, such as the Expenditure Committee (HRSCE 1979, para 68, p23).

Program budgeting was not a significant component of the Committee's original line of inquiry into the budget estimates (although, it should be noted, advisers had envisaged a series of inquiries to follow, one of which would examine the adequacy of processes for evaluating public expenditure (Australian Archives Series Number A 5433. File BE 2a: Papers: Staff and Specialist Advisers. Briefing Paper No. 9)). It was accepted that information on this process was meagre and difficult to assess, not least because of the intermeshing of political and bureaucratic contributions (Australian Archives Series Number A 5433. File BE 2a.). Further, given the recent split of the Treasury Department, the inquiry was also perceived as an opportunity to delineate the respective roles of Treasury and Finance.

In both oral evidence and written submissions, the Department of Finance defined its role generally as providing advice to the Minister on options for policy changes and savings, and briefing on all cabinet submissions requiring expenditure. The specifics, however, were more equivocal. These ranged from assisting cabinet on "how to cut back rationally" and "cutting out the fat" (which reflected the then government's policy priorities of expenditure restraint) to employing an "across the board" approach to formulating the estimates (Australian Archives Series Number A 5433. File BE 2b). At the same time, of course, the Treasury was at great pains to emphasise that it set the parameters in which Finance operated by advising on expenditure trends. While Finance gave considerable credence to the forward estimates process, it also reacted quite coolly to the idea of introducing a program format for the presentation of budget information (note HRSCE 1979, para 72, p24). Given the positive response which came from other central agencies, such as the PSB and DPMC, this came as something of a surprise to both the committee and the secretaries of some operating departments (see, for instance, the response from the Secretary of the Department of Transport. Australian Archives Series Number A 5433. File BE 8.) This reluctance appears to have reflected reservations about introducing new and possibly complicating components to a budgetary process which Finance perceived to be working well, and this despite the fact that Finance had conceded that parts of the same budgetary process already incorporated aspects of "program budgeting".

Bureaucratic intransigence was only one obstacle. The fact that much of the Committee's blueprint was not accepted is not particularly surprising given the apathy displayed by both parliament and the government. The report was tabled in April 1979 and a government response was not forthcoming until March 1980. Further, although the Government accepted the need for program statements and conceded that the specification of objectives would be a "gradual process", it conveniently sidestepped the *rationale* for them—measuring program effectiveness (Eric Robinson, Minister for Finance, CPD, House of Representatives, 4 March 1980, p584).

This, however, did not curtail parliament's interest in the emerging relationship between program evaluation and accountability. Much of the vigour displayed by the Senate estimates committees discussed earlier was transferred to the Senate Standing Committee on Social Welfare through the zeal of its chairman, Senator Peter Baume, and materialised in the form of its report, *Through A Glass Darkly: Evaluation in Australian Health and Welfare Services* (SSCSW 1979). It is a seminal document directed squarely at the role of program evaluation in Australian government administration. Although both it and the HRSCE touched on several points, most notably the need for program statements and developing a role for parliament, the SSCSW followed an alternative route from that of its lower house kin by pursuing evaluation as principally a tool for program management.
Provocatively, from page one, the Report saw:

> in health and welfare in Australia a system out of control—part of a larger crisis in administration; certainly out of the control of the individuals it is supposed to serve and of the institutions and political agencies to which we look for national management. It is also probably out of control of the public servants immediately responsible for its management and of the agencies actually delivering the services (SSCSW 1979, p1).

This assessment set a revelatory tone for the committee's report. Ostensibly the committee was focusing on health and welfare services, but its observations cast a service-wide net, respecifying the need for evaluation activity as fundamentally a problem of accountability for the appropriateness of expenditure and the effectiveness of government programs (SSCSW 1979, pp6, 13, 119). In assessing the level of evaluation activity in Australian government the Committee was "appalled" at the incapacity and apparent unwillingness of health and welfare providers to specify even the broad objectives of their programs, concluding that "evaluation activity in Australia remains inadequate in amount and deficient in quality" (SSCSW 1979, p17).

[1] The evaluation that was undertaken focused on process and structure rather than effectiveness, and consistently failed to ask fundamental questions about the relevance of existing policies or the need for new policies.

Like the HRSCE, the Senate Committee maintained that systematic evaluation was not likely to develop in the absence of either a program of oversight by parliament or a program format for expenditure and results. Unlike the House Committee, however, the Senate Committee was more cognisant of both the theoretical and practical obstacles facing effectiveness evaluation by parliament. As a consequence, evaluation was firmly located as an administrative responsibility of the department delivering the program, but combined with a reporting obligation to parliament:

> it is clear to this Committee that the evaluation process ought to be the responsibility of the department providing a program or service, perhaps with access to specialised units in the Commonwealth Public Service to which it can turn for help with specific evaluation exercises. Centralised authorities should have the task of examining such evaluations with a view to ensuring that they are adequately carried out and that there is an analysis of the overall policy ramifications (SSCSW 1979, p46; see also pp39-40).

In doing so, however, it also flatly rejected any notion of a central evaluation function:

> centralised systematic evaluation, either through the budget process, or as a new central function, is neither practical nor possible. It is not practical, because it does not sufficiently involve the operative staff, and thus is ineffective and more intrusive and threatening than necessary. It is not possible, because the resources that would be

---

[1] In reaching this conclusion the Committee disregarded research undertaken on its behalf that indicated that Australian government was "not suffering from a shortfall in the number of evaluations", only that such activity was usually badly coordinated, not targeted effectively, or undertaken at an inadequate level of analysis (Gross 1979a; 1979b, pp80, 87). Here, evaluation includes most *ad hoc* mechanisms of review: royal commissions, parliamentary inquiries, boards of inquiry and inter-departmental committees. Review activity within departments was just as busy and equally mis-directed, as the submissions from the Departments of Health and Social Security illustrate (Australian Archives Series Number AA 1982/568).

> required would be vast and remote from the sources of data (SSCSW 1979, p46).

These were not the sole reasons. In particular, the Committee was sceptical about the interest shown by Finance and Treasury in evaluation which the Committee feared was confined simply to budgetary functions rather than with the implications of the effectiveness of programs. Further, the "broad brush" approach of Finance in advising on resource allocation concerned the Committee because it was often founded on "poor" quality information (SSCSW 1979, pp17, 40-1). Not surprisingly DPMC was again charged with a coordinating responsibility for evaluation.

The report's anxiety about centralised evaluation is significant, since it indicates that the Committee was either not aware of the RCAGA's reform proposals or disregarded them as fanciful. Of course it also reflects the overriding welfare interests of the committee and its chairman. The committee was definitely aware of the Expenditure Committee's line of inquiry, having suggested that it pursue with vigour the question of program presentation of budget information (Australian Archives Series Number A 5433. File BE 12; see also JCPA 1979, pp42-3), but in contrast the SSCSW highlighted the social perspective on policy evaluation. There were, according to Chairman Peter Baume, two principal aims of the inquiry (Interview 04). The first was to clarify common misunderstandings about evaluation held particularly by parliamentarians. This required breaking the hold of the "no policy issues only resource issues" attitude of parliamentary scrutiny but, at the same time, recognising that the scrutiny role of parliament should not be overplayed.[2] Second, and more importantly, the Committee wanted to "take evaluation into the public service". Baume was confident that once bureaucrats realised that they were responsible for evaluating the results of programs delivered they would act accordingly—"program managers must accept that they have a social obligation to evaluate their services" (SSCSW 1979, p113; see also Baume 1991, pp35-37).

---

[2] Some evidence of a shift can be found in a parliamentary seminar on government expenditure and accountability arranged by the Joint Committee of Public Accounts in 1980. One of the themes to emerge from discussion was that parliament, through its committees, was, for the first time, asking "fundamental questions concerning the 'hows' and 'whys'" of policy. Further, those committees should be encouraging departments to develop a preparedness to question the relevance of programs, especially long-established programs (see JCPA 1981, p113).

Despite the drive provided by Baume, the shift to an effectiveness orientation in Australian government during the late 1970s can be described, at best, as leisurely. The government's response to *Through a Glass, Darkly* provides ample evidence of this. Program evaluation, it was agreed, was necessary, but more so for reassuring the government that "misspending of public funds is not occurring" (Minister for Social Security, Senator Guilfoyle, CPD, Senate, 6 November 1979, p1894). Despite pleadings to the contrary (Senator Baume, CPD, Senate, 6 November 1979, pp1900-01) effectiveness had not yet dislodged efficiency and economy from the apex of thinking in government.

The four years succeeding the *Report* of the RCAGA can quite properly be characterised as a period of institutional jostling and intensive questioning of alternative designs for an evaluation policy. A number of constant themes present themselves as pertinent for any examination of the current expectations for program evaluation, especially the need for program objectives and measures of program performance. What did change throughout this period was the overlapping complexion given these concerns. The RCAGA introduced 'accountable management' to Australia and this undoubtedly instructed the administrative reforms of the early to mid 1980s. Within this frame of reference, both the RCAGA and the HRSCE concentrated on the budgetary and resource allocation process as the vehicle for promoting evaluation—it was part of the central coordination mechanism. Against a backdrop of the perceived decline in parliament's scrutiny role, the HRSCE and the SSCSW initiated parliamentary claims to program evaluation as an accountability tool, with accountability just starting to recognise 'results'. In the end, however, these claims were tempered by institutional realities, the final muted message being "[d]on't do the work yourself, but see that it is done and examine it" (JCPA 1979, p43). As if to lead by demonstration, the SSCSW was the first body to highlight evaluation as part of the policy development and service delivery obligations of public service program managers.

As the tributaries joined the policy stream, markedly different—and partly irreconcilable—expectations for evaluation were emerging. At one level effectiveness reviews as a central instrument of coordination was strongly credentialled by the RCAGA. At another, evaluation was slowly being seen as part of responsible administrative—that is, managerial—practice. As we will see, although the latter was not necessarily incompatible with the former—since evaluation obligations sat comfortably with the increased policy responsibilities of officials under accountable management—in practice meeting the information requirements of both would present itself as a major design problem.

Maintaining Momentum

As the previous section has described, the perceived decline and subsequent revival of parliamentary scrutiny of the executive was an issue which permeated thinking on Australian government during the 1970s, even though the RCAGA did not judge itself competent to provide detailed advice in this area (RCAGA 1976, pp107-123). The type of parliamentary investigations outlined were certainly not occurring in isolation, since this was a period which saw government administration increasingly illuminated by external review, even if it was not necessarily of the effectiveness variety. Two developments in particular sufficiently reflect this change: the new administrative law and the introduction of efficiency audits.

Since the early 1960s there had been a growing number of calls (emanating mainly from the legal profession) for the reform of procedures available to aggrieved citizens for the review of administrative decisions and, as a corollary, the prising open of government decision making. These were orchestrated by the 1971 Report of the Kerr Committee on Administrative Review and follow-up reports. In addition to the procedural complexity and excessive cost of legal remedies for a limited range of maladministration and wrong decisions, the chief criticism levelled by the Kerr Committee was at the patent inability of parliament to ensure adequate scrutiny of the administration on such matters (Goldring 1981, pp84-5; Jinks 1982, pp211, 215-6; Wilenski 1986, pp186-7). Its "extremely far-reaching" remedy was to look beyond parliament by recommending the creation of a system of both internal but mostly external administrative review based on enhanced independent statutory tribunals (for administrative appeals on merits) and investigative officers (such as an ombudsman), the augmentation of judicial review of questions of law and better access to government information (Goldring 1981, pp86-94). It was an area which received remarkable bi-partisan political support and, with the notable exception of freedom of information legislation, was acted on with relative expedition during the mid to late 1970s.[3]

---

[3] The relevant Commonwealth legislation is as follows: Administrative Appeals Tribunal Act 1975; Ombudsman Act 1976; and the Administrative Decisions (Judicial Review) Act 1977 (although this last act only came into operation in late 1980). Although a bill for the final part of the review system—freedom of access to

At around the same time perhaps the only major success of the RCAGA was coming to fruition.[4] Having accepted "in principle" the RCAGA's recommendations on efficiency audits soon after it reported in 1976, the Fraser Government quickly appointed a working group from the public service to examine alternative models for an efficiency auditing function for the Auditor-General. Based on the report of the working group, in late 1978 the government finally introduced a bill to amend the venerable Audit Act 1901 so as to enable the Auditor-General to undertake efficiency audits aimed at assessing both departmental economy and efficiency in the use of resources and procedures for ensuring as much. Packaged as a long overdue necessity for parliamentary scrutiny of the executive, the measures were passed in 1979 (CPD, House of Representatives, 25 October 1978, pp2297-99, 2303).[5]

The introduction of efficiency auditing is particularly instructive because it spotlights the slippery distinction between questions of efficiency and effectiveness. A case in point is the role of the Auditor-General. Primary involvement in reviewing effectiveness and policy was rejected outright by the amendment, but an eminently more realistic secondary role did receive some attention. This could be in the form of either a monitoring function of departmental procedures for evaluating program effectiveness or, in the case of the working group on efficiency audits, explicitly countenancing an informal advisory role for the Auditor-General as part of its efficiency audit responsibilities. This would involve informing coordinating departments such as DPMC on priorities for effectiveness reviews (CPD, House of Representatives, 7 November 1977, p2966; 25 October 1978, pp2298, 2304). Although rejecting as "inappropriate" the extension of efficiency audit powers into policy review—although curiously for the reason that "any findings related to policy would be incidental and therefore incomplete"—this proposed advisory role was affirmed by the Auditor-General in several fora (Auditor-

---

information—was introduced into parliament in 1978 it was only after protracted committee examination and debate that it came into operation in late 1982; see Freedom of Information Act 1982.

[4] A report on the implementation of RCAGA recommendations was compiled by the government (CPD, Senate, 23 May 1980, pp2863-2870). It makes for depressing reading. Among the recommendations rejected by the Fraser Government were the crucial proposals regarding the forward estimates. It was feared that published projections would precipitate "budget rigidity and trauma in trying to alter expenditure patterns" (CPD, House of Representatives, 4 March 1980, p583).

[5] See the Audit Act Amendment Act 1979, s40.

General's submission to the SSCSW October 1978, pp2696-7. Australian Archives Series Number AA 1982/568). This last proposal, which was embodied within the amendment bill as a requirement of the Auditor-General to report, in the first instance, "certain matters" to the attention of Ministers rather than parliament, was criticised as further undermining the scrutiny role of parliament by unnecessarily eroding the parliament-audit office nexus (CPD, House of Representatives, 25 October 1978, pp2304, 2309). The first proposal would be acted on, but not before a decade had elapsed.

Of course, one noticeable feature of the dialogue on evaluation and review in the late 1970s was the taciturn responses coming from government departments themselves. To be sure, the central agencies, and especially the PSB, were enthusiastic contributors to the RCAGA's mission and likewise the Departments of Health and Social Security to the SSCSW. Further, some departments had quietly begun to build up an internal evaluation capacity. For instance, the Department of Health had, since early 1974, established an evaluation section in their Policy and Planning Branch that primarily advised on techniques for the evaluation of hospital and community health services (Submission from the Commonwealth Department of Health to the SSCSW, paras 14-16: Australian Archives Series Number AA1982/568). However, despite these developments the department also conceded that most health programs were implemented without any planning for evaluation. Other departments engaged in public works, transport and communication had also developed capacities for policy analysis and economic cost-benefit studies in the early 1970s (Weller & Cutt 1976, pp137-8). In all likelihood, however, evaluation capacities in these last mentioned departments were more probably a product of both the technical complexity of the policy areas and a longer tradition of analytical support. The administration, it would seem, remained in the wings as a reactive player.

Not surprisingly then impetus for change was to come from outside the commonwealth government. Two principal sources can be identified. The first derives cumulatively from budgetary developments in the various state governments. The 1970s witnessed an unprecedented level of scutiny of Australian state administrations by independent inquiries. One of the most influential was the Review of New South Wales Government Administration of 1977-1982. Predictably, given that its Commissioner, Peter Wilenski, had been briefly associated with the RCAGA and subscribed to much of its reform agenda, the Review's main budgetary recommendations mirrored that of its federal predecessor. Designed to bolster ministerial control, especially within the state Treasury dominated budget process, the Review recommended the development of rolling forward estimates based on a program format that related costs and outputs in order to assist planning and accountability (RNSWGA 1977, pp47-9; Alaba 1994, pp94-6). This was to be supported by a vastly improved analytical capacity within the Premier's Department to evaluate the effectiveness of existing and proposed polices (RNSWGA 1977, pp19-21, 48). Implementation over the ensuing five years was problematic, fuelled predominantly by strong Treasury-inspired recalcitrance but, by and large, these recommendations reflected the new orthodoxy at the state level (Alaba 1994, pp121-29). Other examples are not hard to find, in particular South Australia which began applying aspects of program performance budgeting to the presentation of estimates in the late 1970s and early 1980s (Strickland 1982, pp114-120).

The second impetus for change came from extra-government policy development during the late 1970s. By 1980 the Australian Labor Party (ALP) had been languishing in opposition for a total of twenty-six of the previous thirty years. Learning from the chequered experience of the brief Whitlam Ministries of 1972-1975, the ALP leadership was of the view that one of the major imperatives facing the party once it regained power was to affect change within the administrative apparatus serving the government, so as to render it more "responsive" to ALP policy initiatives (Hawker 1981, p2; Wilenski 1986, p190; Warhurst 1988, pp333-336). Reforms were anticipated in areas as diverse as minister-bureaucrat relations, recruitment and tenure, and industrial democracy, as well as in budgetary decision making and financial management. In 1979 and 1981 the ALP caucus established two separate task forces to examine this brief and draw up plans for government. The end product of these reviews was the policy document *Labor and Quality of Government* which committed a Federal Labor Government to a number of the proposals which had been floating around since the RCAGA reported in 1976. The document was true to its origins, restating the ALP imperative for a public service "responsive to both governmental priorities and to changing community values and needs" (ALP 1983, p14). To this end, budgetary reform was of considerable import and the document pledged the creation of a task force to consider "whether and in what form program budgeting should be introduced". To complement this approach "a system of strategy and effectiveness reviews were to be initiated and efficiency audits expanded" (ALP 1983, p19).

At much the same time, evaluation was also being pushed from other sources. Parliament, for one, had not yielded on the evaluation front, although its point of entry had shifted. In 1982 the Joint Committee of Public Accounts (JCPA) released its report on *The Selection and Development of Senior Managers in the Commonwealth Public Service.* Investigating the prospects for appraising the performance of senior managers in the Australian Public Service (APS), the JCPA followed one of the lines of inquiry pursued by the SSCSW—that "managing" in the public service must involve the initiation, guidance and evaluation of policy (JCPA 1982, para 1.27, p7).[6] Given these attributes the public service could do no better than to embrace "management". More importantly, however, although the committee noted that the "tools of analysis, particularly for evaluating effectiveness in achieving political objectives, are not refined" the development of mechanisms for assessing the effectiveness of senior managers should continue apace (JCPA 1982, para 1.33, p8).

Similarly, the Review of Commonwealth Administration (the Reid Review), which had been established in late 1982 as a response to perceived "administrative breakdown", reported in the dying days of the Fraser Government and also targeted the improvement of management capacities within the APS (Thompson 1989, p218; Dickenson 1984, p44). Like the RCAGA the Reid Review saw considerable merit in proposals to delegate more decision making authority to public officials and saw review as having a supporting role. But in stark contrast to its antecedents, the Review concluded that the level of review (especially external legal) activity in departments was generally excessive—having the result of "diverting the attention of administrators from their broader functional responsibilities"—and should be duly moderated (RCA 1983, paras 5.9-11, p44). Effectiveness reviews, however, were a special case. Since they dealt with "policy-oriented" issues they "properly rested" with the Minister and his or her department (and, in the main, not central agencies) (RCA 1983, para 5.14, p44; para 6.14, p55). As a consequence, Reid provided a perspective on program evaluation which emphasised its role in departmental "policy implementation":

> A well-run department will, as part of its overall internal review activities, look to a variety of mechanisms for ensuring effective delivery of its services. This will include constant questioning, at times by detailed study, of the very fundamentals—whether policy objectives are being met in the best possible way or, indeed, whether the

[6] Wilenski notes that the report was "heavily influenced by individuals associated with the Labor Party Task Force" (1986, p191).

> objectives themselves are appropriate. Whatever the mechanism chosen, the quality of policy examinations of this type will do much to enhance or mar the overall achievement of the goals of government (RCA 1983, para 5.15, p45).

This was also one indication that the SSCSW's version of program evaluation was gaining adherents. Unlike the RCAGA, upon which it also drew, most of the Reid Review's recommendations were enthusiastically embraced by the government and partially acceded to by the Labor opposition. This, combined with the JCPA efforts, helped produce a bi-partisan political climate in the early and mid 1980s receptive to administrative reform (Thompson 1989, p218; Wilenski 1986, p191).

LABOR AND THE FINANCIAL MANAGEMENT IMPROVEMENT PROGRAM

The period spanning from 1976 to the beginning of 1983 was characterised by a great deal of posturing on the issue of evaluating program effectiveness by a number of institutional actors within Australian government. A policy impasse had been reached. Obviously, obstinacy on the part of the Fraser Government was the most serious impediment. Despite its record on the new administrative law and efficiency audits, this government tended to push to the margins reforms to the core administrative machinery. Strained relations between the government and the bureaucracy during this period were a function of Fraser's increasingly blunt attempts to exert ministerial dominance and ensure public service responsiveness (Thompson 1989, pp213-19). Arguably, however, this impasse was also aided by a general misapprehension as to the scope of the new effectiveness criterion for government performance, a task made all the more difficult because of the competing intentions for what purpose evaluation should primarily serve.

What was required by the early 1980s was the political resolve to implement the main aspects of a diverse reform program which had gradually built up a head of steam. The stimulus came in the form of an adroitly prepared Australian

Labor Party which won office in March 1983.[7] The Hawke Government moved quickly to capitalise on its detailed plan for administrative reform. This final section is an overview of the initial reform period between 1984 and 1986. It does not entail a detailed account of all the comprehensive change, but keeping in mind the major trends emanating from the earlier design period, selectively addresses those reforms which have had the most bearing on the development of program evaluation.

Despite some early drag and under the guidance of John Dawkins, the Minister for Finance who had also been assigned responsibility for public service matters, the reform agenda mapped out in *Labor and Quality of Government* was put in train between December 1983 and April 1984 with the release of two Government white papers (Wilenski 1986, pp191-3; Dawkins 1985, pp64-6). The first of these, *Reforming the Australian Public Service* drew heavily on the output of the gestation period of the late 1970s (Commonwealth of Australia 1983, piv). From a wide ambit of concerns the paper focused on two broad objectives: first, making administrators more responsive and accountable to Ministers and Parliament by reinforcing the principle of ministerial responsibility; and second, ensuring that the APS was both more efficient and effective (Commonwealth of Australia 1983, pp1-2; Dawkins 1985, p61). The first was to be secured by allowing for the regular meeting of Ministers so as “to allow a collegiate examination of national prospects and Government priorities” (Commonwealth of Australia 1983, para 3.2.2, p27).[8] Resource allocation figured prominently in efforts to augment ministerial direction:

> The Government is determined that strategy review [at the Cabinet level] and resource allocation processes should reduce the burden on Ministers and simplify their tasks (Commonwealth of Australia 1983, para 3.5.5, p31).

---

[7] Indeed, due to the ALP’s unprecedented levels of preparation, Nethercote suggests that the 1983 federal election “was the first in which administrative policy had claims to be a major issue” (Nethercote 1984, p19; on ALP preparation see Weller 1983).

[8] Other, and arguably just as significant changes, included alterations to procedures for the appointment and role of departmental secretaries which defined the relationship between minister and secretary in “unequivocal terms” and reforms to the structure of the senior executive (Halligan 1988, pp38-45; Commonwealth of Australia 1983, pp10-26). These were cemented by the Public Service Reform Act 1984 (see Public Service Act 1922, ss25 (2) & 26AA: 49-49E).

The method for achieving this was by now quite familiar. The introduction of program budgeting—designed so as to permit an increased ministerial role in departmental resource allocation—supported by effectiveness reviews, the latter explicitly being the "continuing responsibility of program managers" (Commonwealth of Australia 1983, paras 3.4.5 & 3.5.1, p30). Going some way further, however, the white paper envisaged an

> array of program review mechanisms, each appropriate for different purposes and differing with respect to their emphasis on policy, administration and expenditure reduction (Commonwealth of Australia 1983, para 3.5.1, p30).

An environment conducive to this practice was to be achieved by a kind of administrative 'terra-forming' by implementing a financial management improvement program as first recommended by the Reid Review (Commonwealth of Australia 1983, para 3.4.3, p30).

The second white paper, *Budget Reform* sharpened the focus on ministerial direction by setting out refinements to the forward estimates process and Cabinet decision making. In order to improve budgetary decision making the white paper revisited the coordination machinery canvassed by the RCAGA. Of critical importance was the expenditure review committee (ERC) of Cabinet established in March 1983 which, it was envisaged, would have a role in initiating assessments of the effectiveness of continuing programs.[9] The ERC would form the main link between review activity and budget decision making (Commonwealth of Australia 1984, para 2.14, pp7-8). At the same time, by focusing on program effectiveness rather than economy, it was anticipated that program budgeting could provide a "rationally based framework for government decision making, program management and parliamentary and

[9] This particular function of the ERC—the re-examination of existing programs—was probably well established by the time of the second white paper as a response to the adverse fiscal situation which faced the incoming Labor Government (see fn 10 below). The then Prime Minister Bob Hawke recalls that, from the beginning, the role of Cabinet and, one presumes, the ERC was "to clear the arteries blocked by old and misdirected spending programs to make way for fresh initiatives protective of those most in need " (Hawke 1994, p177). The same viewpoint was taken by John Dawkins, the then Minister for Finance: "For a reformist government in particular each budget should be freed as far as possible from history. No program should be sacrosanct and no pattern of expenditure immutable" (Dawkins 1985, p66).

public scrutiny of government activities” (Commonwealth of Australia 1984, para 4.4, p14).

Within this framework the malleability of program evaluation was a virtue which was taken for granted. However, within the *Budget Reform* paper this pliancy was not so well recognised. It countenanced a more rigid design for effectiveness reviews than its companion paper, specifying one of its aims as establishing

> machinery to ensure that the effectiveness and efficiency of programs are reviewed regularly and that the results of such reviews are taken into account in the ongoing evaluation of budgetary priorities (Commonwealth of Australia 1984, para 1.6, p.1).

In June 1983, following Government endorsement of *Reforming the Australian Public Service*, the Department of Finance and the Public Service Board initiated a four month diagnostic study of the prospects for introducing a financial management improvement program (FMIP) as the vehicle for these reforms. This study was instrumental in signalling the shift in Australian government from ‘administration’ to ‘management’ and from compliance issues to performance control. In accordance with the general tenor of reform the diagnostic study reviewed management practices by focusing on “financial and resource use”. Evaluation was seen as a “tool” for greater program efficiency (PSB & DoF 1984, pxi).

Although performance measurement and program evaluation were lauded as accepted practice in sound management, the diagnostic study makes for interesting reading precisely because it overlooks the idea of “regular” evaluation. In its view,

> [w]here evaluations of the overall status and value of programs/projects are undertaken, these evaluations [should] be, where possible, joint line/central agency reviews conducted as a *special exercise* and be oriented to be of value to Government in making choices between ongoing and new programs/projects within the Budget context (PSB & DoF 1984, p37, emphasis added).

To view evaluation as a “special exercise” indicated the *ad hoc* nature of a prescription that, remarkably, countered the main rationale for program evaluation—*systematised* review of programs—and was effectively ignored by the white paper on *Budget Reform*. In all likelihood this unexpected departure reflected the operational difficulties which were thrown up by the study, such as the extent to which realistic objectives could be devised that would satisfy both operational *and* political goals, or that, according to the managers surveyed, the primary function of program management was to run programs and not evaluate them—performance monitoring would satisfy those requirements (PSB & DoF 1984, p11).

The FMIP formally commenced in April 1984 under the auspices of the Department of Finance and the Public Service Board. The logic of the program has been well documented elsewhere (Halligan 1988, pp51-5; Holmes and Keating 1990, pp173-79). Suffice it to say that this logic revolved around the budgetary prescriptions contained in the accountable management model first developed by the RCAGA. The first of these was a renewed aggregate control framework based around a three year forward estimates budget process. The estimates provided a record of the level and composition of expenditure authorised by Cabinet—but determined by the ERC—for the following years subject to no policy alteration. The Hawke Government's first departure from established practice was to publish the forward estimates for the budget year 1983-84, as much for political expedience as principled reform.[10] This was to have a desirable effect, ensuring that the forward estimates "have had to be taken seriously, and this has underpinned their progressive upgrading" (Keating 1990, pp2-3). Gradually the Department of Finance moved to preparing the estimates as an authoritative baseline without collecting bids from spending departments and subject only to changing economic circumstances and policy proposals authorised by Cabinet (Keating 1990, pp1-3; Keating & Holmes 1990, p171).

The second component was a program framework which sought to balance incentive and accountability aspects for administrative activities. The primary incentive was that managerial authority—ideally budgetary authority—should be devolved as far down in agencies as possible; that is, "letting the managers manage". The chief accountability mechanism was that, as a *quid pro quo* for this increased discretion, managers were required to demonstrate improved

[10] Upon entering office the Labor Government had inherited from its predecessor a projected—and, publicly at least, unexpected—budget deficit of some $9.6 billion for the fiscal year 1983-84 (Walsh 1995, pp102-3, 111-115). In order to fully expose both the magnitude of the deficit and the economic 'mis-management' of the Coalition Government, good politics and Labor Party policy of publishing the forward estimates conveniently converged (Kelly 1994, pp56-7, 71; Howard 1986, p61).

program performance not only in terms of process (compliance) but also in terms of outcomes (effectiveness); that is, “making the managers manage”. Program evaluation and techniques of performance measurement were the management tools which would facilitate this. However, for the larger part of this initial reform period they maintained a relatively low profile. Instead the central agencies responsible for coordinating the FMIP concentrated on laying a stable foundation by developing program budgeting.

Program budgeting has been isolated as the “centre-piece” of the FMIP (Halligan 1988, p53). It has been consistently portrayed as the crucial link between the objectives of programs, the measurement of their effectiveness and the determination of priorities in resource allocation. Program budgeting involved the grouping of activities into a program structure according to common objectives, the expression of objectives in measurable terms, an estimation of program costs (which were now to include financial and administrative/staff costs), management reporting systems based on the programs, and a process of review of program efficiency and effectiveness (DoF 1985, p78). The purpose of program budgeting then, was to provide

> a commonly accepted information framework to assist those involved in the public expenditure process to assess [both] program effectiveness and where resources would be best allocated (DoF 1987a, p17).

Thus the intention of program budgeting was true to its name—to allow for more considered budgeting on the basis of programs comparable in terms of total cost and results; that is ‘value for money’. It also indicates why the central agencies believed that the initial emphasis of the FMIP ought to have been directed at developing program structures (DoF 1986a, p19; PSB & DoF 1986, p12; cf Halligan, Beckett & Earnshaw 1992, p27). Difficult as the identification of discrete programs was, it would have been considerably more difficult to attempt to assess activities sprawled, perhaps inconsistently, over whole departments and portfolios. As a consequence, in October 1985 the government announced that program budgeting would be introduced in all departments (DoF 1986a, p16).

The complexion of the reform program was to alter in late 1986 when further extensions were announced in response to Australia’s rapidly deteriorating economic circumstances. Following the deregulation of the Australian exchange rate in late 1983, and particularly since early 1985, the value of the Australian dollar had depreciated alarmingly against the currencies of our major trading partners. Between June 1985 and December 1986 Australia's terms of trade—the ratio of export to import prices—also declined sharply. As a consequence, and in the wake of the Treasurer's caution about Australia sliding into 'banana republic' status, the country's underlying balance of payments problem—and with it foreign debt—ballooned (Gruen & Grattan 1993, pp101-2; Kelly 1994, pp196-199, 202-27; Walsh 1991b, pp40-2). The immediate macroeconomic response was to tighten fiscal policy. This materialised as a forced policy of expenditure cutting and a sharpening of the role for Cabinet's ERC (Hawke 1994, pp377-79). Accordingly, real Commonwealth budget outlays were reduced significantly in 1985-86 and kept constant in 1986-87 (Kelly 1994, p221).

The accompanying microeconomic response was a commitment to increase the competitiveness of the Australian economy. At its base, microeconomic reform is "primarily concerned with the efficiency of production and allocation of goods and services" (Forsyth 1992, pp5, 11-13). At a more specific level it refers to the refashioning of those institutions that form the framework of the economy. The balance of payments crisis exposed structural weaknesses in the Australian economy that, in the eyes of government policy makers, required comprehensive institutional reform. In particular, increasing the efficiency of infrastructure, such as transport, communications and the labour market, predominantly through deregulation, emerged as a priority.

The public service, of course, could not be excluded from both macro and microeconomic reform initiatives. In fact, it represented the most accessible lever of reform. In announcing administrative changes in September 1986 the Prime Minister stressed that further improvements in public sector efficiency

> are an essential part of the restructuring of the Australian economy which has been made necessary by the decline of international commodity prices and their consequent effects on our economic circumstances. That restructuring must involve all sectors of the Australian economy including the public sector . . . A vital ingredient of this process is that Public Service management must also improve (CPD, House of Representatives, 25 September 1986, p1448).

The public service contribution to restructuring was equated solely with an efficiency imperative, so that "the Government will be making less demand on the public purse, which in turn means that the taxpayers will be saving more" (Ibid, p1449). Three measures reflected this renewed economic rectitude. First, departments and agencies were required to achieve yearly reductions in administrative running costs—“efficiency dividends”—as part return for improved managerial flexibility. The use of these savings would be determined by the government. Second, arrangements for the redeployment or retrenchment of inefficient or surplus staff would be streamlined, in particular removing immunity from dismissal. Third, an efficiency scrutiny unit (ESU) was formed within DPMC under the direction of Mr David Block, a senior private sector consultant. The ESU reported directly to the Prime Minister and the ERC of Cabinet. Modelled on the Rayner Efficiency Unit that had been so effectively utilised in the United Kingdom since 1979, its primary purpose was to examine, in conjunction with the relevant departments, administrative procedures, pinpoint inefficient activities and recommend possible savings (DPMC 1987, p35). Confined as it was to matters of administration (SSCFPA 1989, para 2.12, p9; para 2.34, p16), for the two years it operated the ESU provided an almost unparalleled “efficiency” focus which acted as a surrogate for effectiveness evaluation.

At much the same time the PSB was initiating its own form of evaluation in the guise of Program Management Performance Reviews (PMPRs) which were intended to replace the venerable JMRs of the early 1970s. Spurred by the tight resource environment, the government decided to play all its central agency control cards. The Board was instructed to extend its JMRs beyond “traditional concerns with administrative efficiency” to consider how well defined program objectives were in relation to government objectives (PSB 1985, p77). The new PMPRs were intended to complement the FMIP focus on outcomes. However, despite the likelihood that they might impinge on policy areas PMPRs were not to become embroiled in the “evaluation of policy per se” (PSB 1985, p79). It is doubtful that these reviews were able to achieve very much, especially given the bind the PSB placed itself in with regard to evaluating “policy” and the *ad hoc* nature of the requests for PMPR assistance noted by the Board (PSB 1985, p83). In any event, the PMPR concept was short-lived on account of the demise of the PSB in July 1987.

Although evaluation-like activity was being coordinated from several of the central agencies during the period 1984-1986, it was precisely this fragmentation and relative lack of sophistication (to borrow Cutt’s term) that precipitated the eventual placement of responsibility for evaluation policy with Finance. Indeed, it was during the heady days of restructuring in 1986-87 that Finance assumed a greater role in coordinating the FMIP. As a corollary, the link between public management, evaluation and resource use was tightened (Halligan 1988, p34).

Through the FMIP the government had certainly given a clear commitment to a system of service-wide program evaluation which was to feed into resource allocation decisions at both the central and departmental levels. Progress, as ever, was slow. Reporting on the status of the FMIP in mid 1986 the PSB and Finance presented a patchy record in confirming this. For instance, less than one quarter of Commonwealth agencies had established a “formal policy” for evaluating the performance of their programs, understandably delayed by the emphasis on the cautious development of program structures (PSB & DoF 1986, pp12, 18). This ‘snapshot’ of progress reflected the low, almost non-existent profile which evaluation had as well as the absence of formal guidelines, at least until mid 1986 when Finance reported that draft guidelines on the management of program evaluation were to be circulated to departments in July 1986 (DoF 1986a, p63). It also provides some indication of either the difficulty some departments—particularly those in policy advising and regulatory areas—were encountering in approaching the task of evaluation and review or, just as likely, the reticence of others more sceptical of its benefits (PSB & DoF 1986, pp33-5).

CONCLUSION: EVALUATION IN THE POLICY STREAM

Using the RCAGA as a starting point, this chapter has followed the various tributaries, of varying speed and strength, flowing into the evaluation policy stream. It has indicated that these tributaries have each deposited different layers of expectations for an evaluation policy. As part of its wide-ranging consideration of government administration, the RCAGA introduced the organising concept of ‘accountable management’ to Australia. Within this framework program evaluation was intended to serve a priority setting and budgetary function in order to strengthen political (ministerial) control of both policy and administration. For accountability purposes a distinction was made between efficiency and effectiveness, with external (parliamentary) accountability confined to questions of efficiency. It followed that evaluation was essentially an instrument for central agency policy coordination.

At much the same time, however, governments in Australia were exposed to a climate of increasing external review that placed a premium on issues of public accountability. As a response to perceived deficiencies in parliamentary oversight, a set of two parliamentary committees seized on program evaluation and effectiveness review as the key mechanism of *external* accountability for the executive. Evaluation was linked to the provision of refined performance information through program budgeting, but unlike the RCAGA program effectiveness was portrayed as the chief accountability criterion, although this differed according to the orientation of the committee. For the HRSCE, evaluation could be a parliamentary function. As a consequence, the relevance of policy was not so important as a capacity to ascertain the extent to which programs were achieving objectives. In order to preserve parliamentary committee oversight political sensitivities were recognised. In contrast, the SSCSW was more cautious on overt parliamentary involvement in evaluation, preferring instead a supervisory role in relation to the administration. Evaluation was first and foremost the responsibility of the department managing the program. This did not, however, lessen the emphasis on accountability; departments were expected to demonstrate the appropriateness and effectiveness of government programs not only to their immediate political masters but also to parliament and the public.

This parliamentary interest also signalled a new focus on evaluation as part and parcel of good program management. Both the SSCSW and the Reid Review firmly located the evaluation function within individual departments. This appeared to reflect concerns with program improvement and service delivery considerations rather than purely resource allocation or policy coordination. As a consequence a stand-alone evaluation role for central agencies, and particularly the budget agency, was unequivocally rejected by these bodies.

The various tributaries to the evaluation policy stream indicate what is perhaps the major complication facing the development of program evaluation within government, especially parliamentary government—what, precisely, is its primary function? Not that this problem is confined to Australian government. Comparative analysis has confirmed that, relative to the traditional functions within government such as auditing and budgeting, program evaluation “sometimes struggles to establish clearly whom (or what) it is supposed to serve” (Gray & Jenkins 1993, p190; Dobell & Zussman 1981, pp406-411). The widely diverging complexions and interpretations given to program evaluation in the first phase of development in Australia bears this assessment out remarkably well. Was evaluation primarily an executive tool for assisting central policy coordination and priority setting? Or should it be directed at the policy implementation level so as to increase program responsiveness to community needs? Alternatively, perhaps evaluation is primarily an external accountability mechanism, directed at increasing the transparency of government stewardship of public resources?

While many tools of government serve more than one purpose, there is usually agreement on a primary purpose that shapes the limits of secondary functions. A good example is the external audit function which is closely associated with relatively narrow constitutional principles of accountability and probity. External audit is primarily concerned with questions related to management control and procedures. Even though external auditing has broadened the scope of the questions it asks by examining the efficiency of government operations, this move has been conditioned by the traditional accountability role so that it remains concerned with questions of administration rather than policy effectiveness. In contrast, as this chapter has illustrated, there was no commonly accepted functional role for program evaluation (cf Davis 1990, pp35-37; Rist 1989, pp355-356, 362, 367-369).

What this chapter has shown is that the intended purpose of program evaluation is a function of the institutional perspective of the proponent. Given the absence of clear delineation of the boundaries of effectiveness review, it was not surprising that program evaluation proved to be a fiercely competitive policy arena. In many instances conflicting purposes emerged, pushing and pulling evaluation policy in different directions. Despite the entry of political resolve to the evaluation policy arena in the form of the Hawke Labor Government, there was still a remarkable level of confusion as to where evaluation should fit in. In the end a program evaluation policy would be imposed. Even though the RCAGA tributary proved to be particularly strong, as the next chapter will detail, the eddies and currents caused by institutional tensions ensured that a central evaluation strategy would encounter a number of implementation difficulties.

Chapter Five

The Evaluation Strategy 1987 - 1995

The purpose of this chapter is to complete the story of program evaluation in Australian government. It is a detailed historical account of the development of the current evaluation strategy and its implementation. This is presented as a series of incremental responses to a deteriorating policy environment and demonstrates that the evaluation strategy is unapologetically a budgetary initiative. The strategy, of course, cannot be divorced from the functions of the central agency responsible for it, the Department of Finance, which has, due to the policy imperatives of the Labor Government, established itself as one of the most influential agencies in the Commonwealth bureaucracy. As a consequence, the next chapter will also assess the role of the department and the important relationship between the policy management roles of Finance and program evaluation—that is, budgetary control, policy advising and framework management.

Throughout this period budgetary priorities were the dominant theme. An important consequence of this was the increasingly tight link between the Department of Finance and the expenditure review committee (ERC) of cabinet. Indeed, some have gone so far as to characterise the Department of Finance, or at least its general expenditure division, as the *de facto* secretariat of the ERC (Campbell & Halligan 1992a, p31). To some degree, the accuracy of this description is perhaps most vividly borne out by the development of the government's evaluation strategy, since senior Finance officials, with the strong support of the Minister for Finance, often sought to secure cabinet decisions to give weight to the importance of evaluation in the government's fiscal objectives. This by itself reveals much about the priority accorded evaluation by the Labor Government, since the later period of this administration was characterised by a marked reduction in—and hence focusing of—cabinet activity (Weller 1992, pp6-8).

The dividing line for analysis undertaken in Chapters Four and Five has been set at the end of 1986 for several reasons. In the first place, machinery of government changes in July 1987 saw a level of restructuring that both affected the complexion of the reform program and further elevated economic policy considerations. This spurt of redesign has been identified by its principal architects as the "dominant structural element of reform" (Holmes & Shand 1995, p567). Second, and as a consequence of the first, the role of the Department of Finance as the linchpin of both managerial reform and fiscal restraint was magnified. It was in this period that program evaluation policy was prioritised both by giving Finance responsibility for it and making it mandatory. That is, evaluation policy finally moved from conception to actuality.

ALIGNING ECONOMIC AND ADMINISTRATIVE PRIORITIES

A celebrated chronicler of the Hawke-Keating Labor Government has noted that "virtually every economic milestone in the early years was a response to a crisis" (Kelly 1994, p226). The same can be said of administrative reform. The economic disorientation brought on by the balance of payments crisis of 1986, briefly discussed towards the end of the previous chapter, was ultimately to determine the direction of administrative reform—and the evaluation policy stream—for the remainder of the decade. The third phase of the Hawke Government's administrative reform agenda, announced almost immediately after the July 1987 federal election, was dictated by the ever-sharpening twin economic priorities of the government: budgetary restraint and microeconomic reform. Importantly, not only were these priorities reflected in both new ministerial and administrative arrangements, they were institutionalised in a new configuration for cabinet committees in August 1987 (Codd 1990, pp6-10; Weller 1990, p22).

The first imperative was budgetary restraint. The vulnerability of the Australian economy—and specifically the infirmity of the Australian dollar—was a source of continuing policy and budgetary problems in the mid 1980s. The desperate budget of 1986 illustrates this particularly well. In order to halt a precipitous slide of the Australian dollar, budget deliberations were re-opened after policy had been set and a rigorous process of "slash and burn" adhered to in order to gain expenditure reductions sufficient to appease volatile currency and investment markets (Hawke 1994, pp377-380; Walsh 1995, pp149-156). The driving force behind the search for savings was the 'troika' of economic ministers who, under the chair of the Prime Minister, constituted the formidable standing ERC of cabinet (Kelly 1994, pp273-4). From 1986, under the firm direction of the ERC, the Labor Government was to pursue a budgetary strategy that utilised targets designed to achieve real reductions in Commonwealth outlays. Indeed, in 1987 the Labor Government secured the first Commonwealth budget surplus in thirty years.

The second priority was the acceleration of wide ranging microeconomic reform that encompassed the pursuit of market liberalisation for government business enterprises, deregulation of financial and labour markets and hefty reductions in protection for most manufacturing industries (Kelly 1994, pp390-398; Walsh 1991b, p42). This policy area was closely co-ordinated by a new structural adjustment committee (SAC) of cabinet that, along with the ERC, was judged to have been marked by a high level of ministerial assiduity (Codd 1990, p7; Weller 1990, p22). While the 1987 administrative changes were to reflect these pressures for structural adjustment they were, by and large, a direct descendent of the savings measures of the year before, and confirmed that public service reform had "become oriented more to short-term concerns of government—the immediate political requirements" (Halligan 1987, p47).

On 14 July 1987 the Prime Minister, Mr Hawke, buoyed by a third federal election victory, announced a comprehensive restructuring of the Commonwealth administration (Hawke 1987; CPD, House of Representatives, 15 September 1987, pp43-6). The economic shock wave of 1986 was still reverberating in the rationalisation of the changes. Economic reform was

> not simply a task for our trade-exposed sectors like agriculture, mining or, increasingly, manufacturing and tourism. It is a task for the whole economy, because all sectors can suffer if one is inefficient. In applying measures to deal with the responsible reconstruction of the economy, we are determined to apply the same rigour to the Government's own household (CPD, House of Representatives, 15 September 1987, p43).

Three of these measures are of significance here. The first was the creation of an enlarged two-tiered cabinet structure consisting of sixteen portfolios each of which would be the responsibility of a cabinet minister. The second tier was comprised of non-cabinet or junior ministers who would support the cabinet minister by assuming responsibility for "specific parts" of a portfolio.[1] In order to facilitate these new arrangements the number of departments was drastically reduced from twenty-eight to sixteen, mostly by way of consolidation. This heralded the creation of the so-called 'mega-departments'.

The third change was a fundamental revision of central agency functions which derived from recommendations made by the Government's Efficiency Scrutiny Unit. The most audacious move was to abolish the venerable Public Service Board (PSB) and distribute most of its establishment functions to other departments. In particular, public service arbitration, pay and conditions issues were transferred to the Department of Industrial Relations; operational aspects of personnel management were devolved to individual departments; and both classification matters and responsibility for management improvement and program budgeting were placed with the Department of Finance. A Public Service Commission would take the place of the PSB, having responsibility for residual functions relating to the policy aspects of recruitment, promotion, discipline and retirement (Hawke 1987, p13). The changes received bi-partisan support (CPD, House of Representatives, 15 September 1987, pp46-7).

Two points can be noted here. Although some commentators understandably lamented that the dissolution of the Public Service Board "effectively leaves the Australian public service without a focal point for management and development", and especially in regard to personnel matters, the new administrative arrangements were clearly relocating responsibility for management reform with the Department of Finance (Nethercote 1988, p15; CPD, House of Representatives, 17 September 1987, pp217-18). This augmented role of Finance was to be guided by a new Australian Public Service Management Advisory Board (MAB) that was charged with a purely advisory role. In this capacity it was to provide a "forum for consideration of major management activities affecting the Australian Public Service as a whole" (Public Service Act 1922, ss22(2)(a) & (b)). The MAB comprises the upper echelon of the APS. It is chaired by the secretary of DPMC and includes *ex officio*

---

[1] A portfolio refers to the departmental policy responsibilities of a cabinet minister. It consists of at least one department of state, usually a 'mega-department' that covered several policy areas (for example, employment, education and training). In most cases a cabinet minister delegated specific departmental responsibilities (from the previous example, say employment services and youth affairs) to a junior (non-cabinet) minister. In this way, each department had cabinet level representation through the responsible portfolio minister.

the secretaries of Finance and Industrial Relations, and the Public Service Commissioner as permanent members (Public Service Act 1922, s22(3)). These new administrative structures were consolidating arrangements for central management direction.

Second, it was inevitable that the consolidation of departments would rupture aspects of the reform program. It was readily conceded by the Prime Minister that the “primary impetus” for this reworking was increasing efficiency (Hawke 1987, p13). Efficiency, however, was dressed in a number of guises. The new cabinet structure was intended to “enhance ministerial control of departments” while the amalgamation of departments was justified by a number of subsidiary reasons, but in particular

> to achieve administrative efficiencies and savings, better policy co-ordination and improved budgetary processes (Hawke 1987, p12; 1994, p416; CPD, House of Representatives, 15 September 1987, p46; 17 September 1987, p219).

Perhaps the most crucial of these was the objective of increased policy coordination. The early phase of the FMIP was cautiously directed at laying a solid foundation for budget and resource management improvement in the form of program budgeting. While the creation of mega-departments purported to aid policy coordination by introducing “broader perspectives”, it also severely disrupted the diffusion of program budgeting. It was recognised at the time that, by and large, the creation of portfolios suffered from a lack of clear direction, adhering neither to the organising principles of function or policy field (Halligan 1987, p42; CPD, House of Representatives, 17 September 1987, p201). Subsequent assessments of this period confirmed that several of the mega-departments married unrelated programs, multiple and sometimes conflicting policy objectives and bureaucratic cultures that, as a matter of course, made for a long adjustment period. In many instances, the departmental program structures which had been so carefully cultivated required expeditious reconstruction (Campbell & Halligan 1992a, pp135-6, 177-183; Halligan, Beckett & Earnshaw 1992, pp13-15).[2]

---

[2] The impact of departmental amalgamations on aspects of policy coordination has received considerable attention following the 1987 changes. Participants in the amalgamation process have provided widely divergent assessments. Michael Keating argues that policy coordination has been enhanced and that the amalgamations provided the “institutional framework” for portfolio budgeting, while Peter Walsh,

A precursor to the organisational renovation of mid 1987 can be found in a progress report from the Department of Finance on the FMIP, *FMIP and Program Budgeting: A Study of Implementation in Selected Agencies* (hereafter the *Study*) (DoF 1987b). Completed in March 1987—prior to the Bastille Day pronouncements—this study examined "the impact in selected departments of changes in the budgetary environment as a consequence of the FMIP and program budgeting" (DoF 1987b, p4).[3] It is a significant document for two reasons. First, it provides an insight into the developmental processes of the FMIP and particularly attempts to placate emerging tensions between the centre and line departments over the complexion of performance based budgeting. In doing so it clearly identifies program budgeting as the linchpin of performance budgeting and evaluation under the FMIP. Second, it is an assertion of Finance's newly augmented coordinating role for the FMIP, recommending that Finance assume responsibility for the development of performance measurement (DoF 1987b, pp2-3).

Like its earlier counterpart, the 1986 *FMIP Report* (PSB &DoF 1986), the *Study* articulates "corporate management" as the framework guiding the FMIP. Corporate management consisted of three linked components: corporate planning (which involves defining the organisation's goals and objectives), program budgeting, and management information systems and performance evaluation (DoF 1987b, p32; Weller & Lewis 1989, pp7-8). Its overall conclusion was that progress in implementation had been "satisfactory". While it also confirmed that the main benefits of program budgeting had been predominantly "structural"—mainly because program budgeting had only just been introduced in 1985-86—these "benefits are only beginning to be realised and this mainly in areas where outputs are easily measurable" (DoF 1987b, p1). The constraints on development were quickly identified: a lack of performance information and the environmental condition that budget allocation decisions were being determined by short-term macroeconomic factors.

Minister for Finance 1984-1990 is highly critical of the consolidated portfolios—which he labelled "unwieldy and unbalanced"—and sceptical that any administrative savings were actually achieved (Keating 1993a, pp2-10; Keating 1990, p11; Walsh 1995, pp170-1). More systematic analysis has revealed that mega-departments have had "modest success" in achieving policy coordination objectives (Craswell & Davis 1993, p206).

[3] The sample comprised the Departments of Employment and Industrial Relations, Finance, Immigration and Ethnic Affairs, Primary Industry, and Trade.

The *Study* was more forthright than its predecessors in asserting that the central aim of the FMIP was to "establish performance measurement as the basis of budget allocations" (DoF 1987b, p7). Although there had been considerable change to the budget process in the first few years of the Labor Government—including the refined forward estimates process and intensified cabinet committee (ERC) consideration of budget submissions—the tenor of budgetary decision making, especially during the search for savings in the 1986-87 budget round, was still oriented towards an aggregate expenditure limit. This, the *Study* cautioned,

> brings into question much of the rationale of the budget as an allocative mechanism . . . the incentive to evaluate and justify budget estimates on output-related criteria is greatly weakened if performance data do not consistently and significantly affect the allocations (DoF 1987b, p10).

As a result further budgetary changes were required to "reconcile" results-oriented management with the increasing macroeconomic pressures.

The proposed solution was the introduction of medium term resource agreements between the Department of Finance and individual departments so as to provide much needed stability in resourcing. Resource agreements were also intended to provide Finance with a focus for the joint development with departments of appropriate performance measures (DoF 1987b, pp11-12). The *Study* noted that the development of measures to evaluate government expenditures had been a source of "divergence" between Finance and line departments because of the expectations held by each for the use of performance indicators. The Department of Finance concentrated on indicators for budget resource allocation while line departments saw indicators as assisting internal management. The chief task facing Finance was to develop "a clear relationship between budget allocation and performance" since without this

> there is little incentive for the agencies to examine indicators from the point of view of external evaluation. In the absence of indicators relevant for budget purposes, Supply Divisions are tempted under pressure of the budget timetable to continue to rely implicitly on

> input data and comparisons with previous years (DoF 1987b, p14; see also p52).

Program budgeting provided the key to clarifying this relationship since it was explicitly designed to facilitate strategic resources management, the primary objective being “to aid strategic control rather than day-to-day control” (DoF 1987b, pp44; 52, fn2). Program evaluation provided an analysis of performance which was intended to influence strategic decision making in terms of both policy and resource allocation. According to the *Study* there was, by early 1987, scant evidence to suggest either that departments were using program structures for anything other than presentational purposes or evaluating program effectiveness for the purposes of either policy making or resource allocation (DoF 1987b, pp34, 40-1; see also Barrett 1988, p55). Hence, by emphasising the role of evaluation and review the *Study* is notable because it provides a thinly veiled caution that the use of program evaluation as a budgetary instrument had, by the time of the economic crisis of 1986-87, reached a crossroad. It was effectively arguing that if the Government was going to make performance budgeting work, it must do so within the prevailing economic climate of fiscal constraint. This analysis, aided by the ferment caused by the administrative changes of July 1987, intensified the urgency of developing a program evaluation strategy that was closely linked to the budgetary process.

A Sharpened Focus On Program Evaluation

If there was one overriding objective of reformers in the period following the RCAGA it was to establish a *system* of evaluation whereby all programs were assessed regularly and in which the results were used in decision making across different levels within government. Of course this is not to deny that the Labor Government, like its predecessors, had undertaken a great deal of policy ‘review’ through the trusted mechanism of *ad hoc* government appointed inquiries (Prasser 1988, pp120-134). Clearly, however, by mid 1987 the Government’s objective of systematic evaluation had not been achieved. As the *Study* had shown, external effectiveness evaluation—especially in the form of budget related performance indicators—was perceived by departments as intrusive and costly, and a conspicuous absence of strong central coordination hampered its diffusion. Furthermore, as an approach to public management which called for the use of specific analytical techniques, evaluation remained somewhat remote from program managers who continued to view it as a special skill which had to be "imported" (DoF 1987a, pp1-2).

There were two developments that sharpened the focus on program evaluation. The first was the articulation during 1986-87 of the rationale and appropriate techniques of evaluation relevant to government. The second were changes to central budgetary processes at the 1988 budget round. Again, both changes were a direct response to the overriding objective of fiscal restraint. Each will be considered in turn.

In early 1986, as part of the program budgeting drive, the Department of Finance in conjunction with the Public Service Board sought to correct some of the misgivings outlined earlier by initiating the development of a guide to program evaluation, the general idea being to produce something of use to the evaluation "layman" (DoF 1986a, p63; Interview 10). This was circulated throughout government departments in mid 1986 and finally emerged in May 1987 as a joint publication, *Evaluating Government Programs: A Handbook* (hereafter the *Handbook*). This was the first attempt by central agencies to articulate a "framework" for program evaluation, covering its purpose, orientation and alternative methodologies.[4] Evaluation in government, it was claimed, served a similar role to the market mechanism in indicating "value for money" (PSB & DoF 1987, pv). In so doing Finance and the PSB were careful to tread lightly within the new environment of devolved management. The *Handbook* did not prescribe specific techniques but rather "suggest[ed] principles [for the collection, analysis and presentation of data] which might usefully guide the conduct of evaluation" (PSB & DoF 1987, p1).

---

[4] The genesis of the Commonwealth *Handbook* is likely to have benefited from the early experience of program evaluation reviews in New South Wales. In November 1984, in response to recommendations contained in the Wilenski Review of NSW Government Administration, the NSW Government endorsed a statement on comprehensive management audits of departments that included effectiveness reviews. It also established a Program Evaluation Advisory Committee (PEAC) comprising representatives of the central agencies and, to support it, a Program Evaluation Unit (PEU) within the NSW Public Service Board which began operations in January 1985 (NSWPSB 1985, pp46-7). The role of the PEAC and the PEU was essentially to coordinate specialist staff training in evaluation techniques, develop program evaluation guidelines and standards and ensure that departments develop and submit forward five year plans for effectiveness reviews of their programs. A program evaluation policy—Program Performance Review—was subsequently endorsed by the PEAC in May 1986 (Johnston 1992, p35). As the remainder of this chapter will show—and despite initial implementation problems with the NSW policy (see Johnston 1992, pp36-38)—similar structures and policy designs were utilised by the Commonwealth Department of Finance after 1987.

Program evaluation was defined in rather catholic terms as the "systematic assessment of all or part of the program activities to assist managers and other decision makers" to appraise the continued relevance and priority of program activities; test whether program outcomes achieve the stated objectives; canvass alternative ways of achieving the objectives; and decide on the level of resourcing to the program (PSB & DoF 1987, p3). Its "*principal*" purpose, however, was to assist decision making on the allocation and application of resources (PSB & DoF 1987, pp.v, 1). This view was, in turn, focused more clearly as the *Handbook* identified the primary audience for program evaluation as the agency executive and relevant portfolio Ministers. "Other parties" with an interest might include central agencies, cabinet policy committees and program managers (PSB & DoF 1987, pp9-10). This leaves little doubt that the results of evaluation were to feed into *central* resource allocation decision making processes rather than program management. It is also significant because it makes no mention of parliament—and by extension, the public—as a user, thereby disqualifying that institution's claim to this type of review information.

Program evaluation then, as conceived by the *Handbook,* is an internal tool of executive management, but one of increasing sophistication, since a distinction can be made between strategic and tactical evaluation (PSB & DoF 1987, p46). The former is concerned with long term planning and resource allocation and includes ex-ante evaluation (asking whether there is demonstrable need for the program) and effectiveness evaluation (the extent to which objectives are achieved and whether the objectives are still relevant). Tactical evaluation, on the other hand, is undertaken in support of program operations and includes the more 'conventional' forms of evaluation which examine aspects of process and efficiency. Although there was a high likelihood that the distinction could be "blurred" (PSB & DoF 1987, p49), it would seem that effectiveness was only of import at the strategic level.

The *Handbook* may have provided some direction for departments groping for assistance in the evaluation dark[5] , but the 'systematic' part of its definition of evaluation was unlikely to materialise without some central coordination. As a harbinger of its response to economic and administrative policy imperatives, in

[5] Recollections by some Department of Finance officers were more caustic, suggesting that the *Handbook* was "not particularly user-friendly" (Interview 10).

an April 1987 cabinet decision the Government acknowledged this deficiency and determined that ministers and their departments develop annually three year rolling plans for the systematic evaluation of program efficiency and effectiveness, report the outcomes of evaluations to government and that the Department of Finance monitor progress (FD: 88/0621). In addition, all new policy proposals were to be supported by relevant performance measures and an evaluation strategy for future assessment. Following the policy pause occasioned by the federal election, in September 1987 the Department of Finance circulated to the heads of corporate service divisions in departments broad guidelines for the government's policy on evaluation planning (DoF 1987b). The first evaluation plans were to be lodged with Finance by mid December 1987 (DoF 1987b, para 1, p1). This decision was taken within "the context of considering progress with the implementation of program budgeting" and was intended to "encourage departments to focus on the 'real' management aspects" of program budgeting (DoF 1987b, para 2, p1; FD: 88/0621). A standard format for the plans was not proposed, but it was expected that the first plan would include information on major current evaluations and those completed in the preceding three years.

The second inducement for program evaluation was the introduction of portfolio budgeting by the ERC of cabinet at the 1988 budget round and its subsequent refinement in 1989. As discussed in the previous chapter, the forward estimates process underwent considerable enhancement under the first two Labor Governments that placed the forward consequences of policy and resource decisions in relief. According to the then Secretary of the Department of Finance:

> Forward estimates now provide an accurate guide for the development of budgetary strategy; they serve as a base against which budget decisions can be taken and against which the government's fiscal performance can be assessed (Keating 1990, p3).

From 1986-87 the Government's budgetary strategy was straight forward: the imposition of top-down targets designed to effect a real decline in budget outlays. In the ensuing period the ERC was faced with an onerous workload in relation to the vetting of portfolio spending proposals. Within the strategic limits it had set itself, the ERC employed a number of processes for reviewing policy and, at the same time, reducing the deliberations over minor proposals, of which the most effective was portfolio budgeting. First introduced at the

1988 budget round, portfolio budgeting was intended to provide portfolio ministers with flexibility and incentives to allocate resources within fiscal aggregates. After the ERC set savings targets for each portfolio, Ministers were permitted to introduce new programs only on the condition that they offset them against a saving elsewhere in the portfolio. The achievement of portfolio savings targets—that is, the designation of spending cuts—were, at least nominally, left with the individual minister and portfolio (Keating 1990, pp9-11).

A filtration process was used in 1988 to confine portfolio targets to minor programs. In 1989 the process was altered significantly so that *all* new policy was subject to the portfolio targets. This change was “primarily intended to facilitate the reallocation of resources to accommodate high-priority new policy within the headroom afforded by the government’s fiscal target” (Keating 1990, p11). It also presented itself as a logical extension of the ministerial (that is, portfolio) reorganisation of July 1987. However, of particular relevance for our purposes is the fact that portfolio budgeting obliged portfolios to hone their competencies in prioritising their programs in terms of both efficiency and effectiveness (Barrett 1988, p52; cf DoF 1988b, p90). This, in turn, presupposed a capacity for evaluating programs against these criteria. In order to assist the attainment of pressing macroeconomic objectives, the virtues of budgetary devolution were to be realised through a system of program evaluation.[6] This policy setting was to take on even more significance in the period between late 1988 and late 1991 as the current account deficit problem worsened and the economy moved into a prolonged recession (Kelly 1994, pp487-497; Stutchbury 1990, pp72-74). Most tellingly, as a result of fiscal discipline that had already delivered a surplus, the government faced fewer and fewer budgetary options.

Designing The Evaluation Strategy: The Evaluation Task Force

Although broad guidelines were now in place, the outcomes from the first round of evaluation plans were disappointing. According to the Department of

[6] The efficacy of portfolio budgeting as a tool for achieving these macroeconomic objectives has been questioned by the then Minister for Finance, who argues that despite the good intentions the process was often circumvented by post-budget additions (Walsh 1995, p105). There is also some doubt over the extent of budgetary devolution to portfolio ministers (MAB-MIAC 1993a, pp232-233).

Finance's annual report evaluation plans submitted by line departments were of "variable quality" with a sizeable proportion of departments requiring considerable assistance from Finance officials (DoF 1988a, p13). An internal assessment report prepared by the Program Budgeting Branch was less charitable: "by any measure the standard is unacceptably low" (FD: 88/0621). By the beginning of February 1988, of the eighteen portfolios and departments to which requests had been sent, ten had failed to submit a plan while the remainder had managed to submit "a plan of some description" (FD: 88/0621). The deficiencies were not difficult to anticipate. Of the eight plans submitted most—but noticeably those of all policy departments—were merely lists of evaluation-type activity undertaken within the portfolio that made no effort to project policy and resource requirements for evaluation. In part this response could be explained by the generality of the guidelines issued in September. It was also a function of a period of furious expenditure slashing, when evaluation planning was perceived as yet another non-program burden that had to be met.

For the Department of Finance the cool departmental response to planning initiatives was confirmation of the "high potential costs" of pursuing a formal evaluation policy. Of most concern were the risks of developing a compliance approach without really altering existing operations and centralising program evaluation so as to "alienate" spending departments, as well as the residual fear that evaluation would fail for "lack of political acceptance" (Program Budgeting and Management Branch, February 1988. FD: 88/0621). Even at this late stage Finance officers realised they still faced two tasks—convincing both spending departments *and* the political executive of the benefits of evaluation. A preferred strategy did seem to emerge. It relied on securing support at the political level by framing evaluation as serving the pressing need for "major savings", while at the same time cajoling departments into embracing evaluation as part of their own program management responsibilities (Program Budgeting and Management Branch, February 1988; Division Heads Meeting, April 1988. FD: 88/0621). Several mechanisms for initiating an evaluation policy were canvassed, including the establishment of either a high level advisory committee directed at policy development, task-specific working groups or a specialised unit within Finance (Program Budgeting and Management Branch February 1988. FD: 88/0621).

In mid March 1988 the Secretary of the Department of Finance, Dr Michael Keating, adopted the first proposal by establishing an Evaluation Task Force within the Department of Finance (Minute 16 March 1988. FD: 88/0621). In the same month Keating wrote to all departmental secretaries informing them that the purpose of the task force was to develop a "a strategic approach to program evaluation" (Finance Secretary to Departmental Heads, 23 March 1988. FD: 88/0621). Given the "mixed" departmental responses to planning guidelines, it is clear from the Task Force's terms of reference that its assignment comprised both conciliation and agenda setting functions. On the first, it was expected to consult with individual departments on their needs for program evaluation and incorporate this into its "practical" work of assisting departments with the development of evaluation plans and their capacity to undertake evaluation (Evaluation Task Force 1988, p1; DoF 1988b, p66). No doubt this task would be assisted via the external appointment of Linda Lipp, an experienced branch head from the former Department of Employment and Industrial Relations, to lead the Task Force. At the same time the Task Force was establishing "the final structural element" of program budgeting by developing a comprehensive strategy for the enhancement of evaluation within the Australian Public Service and

> identifying systems which will assist in integrating evaluation into program management, policy and budgetary decision making processes of government and departmental level (Evaluation Task Force Terms of Reference. FD: 88/0621).

Of necessity this last objective required the Task Force to clarify the role of the Department of Finance in evaluation.

In the period during which the Task Force undertook its investigations, senior officers within the Department of Finance immersed themselves in a consideration of the tensions that a centrally directed evaluation policy would throw up and how to either avoid or manage these problems. It was, of course, accepted that the role of the Department of Finance was to act as a "catalyst and guide". However, with some disingenuousness, it was conceded that in order for evaluation to gain currency there "will be pressure to treat it not as a savings exercise but as an enhancement exercise" (Division Heads Meeting, April 1988. FD: 88/0621). A projection of this last variety would necessitate questioning of government programs in the public arena, and so there was uncertainty as to whether this was what the government actually desired. Despite its protracted genesis, concerns were now arising over the pace of change: the question had changed to whether Finance was "going too far too fast with evaluation" and expecting portfolios to grapple with political sensitivities for which they may not be appropriately equipped nor particularly receptive (FMIP Steering Committee Meeting, March 1988. FD: 88/0621). Ultimately, however, evaluation could not be divorced from the budget process and most attention was directed at the mechanics of integrating evaluation plans into the central resource allocation decision making process.

In May a progress report from the Evaluation Task Force identified the core tensions of a centrally inspired evaluation policy along similar lines (FD: 88/0621). In the main, departments displayed a deep seated cynicism about evaluation, claiming that many programs were simply not evaluable (especially those relating to policy advising) and that the evaluation drive was, at its base, a "cost-cutting" task and hence inimical to departmental interests. This view was reinforced by an antagonism towards the idea of Finance—or any central agency—having involvement in program evaluation planning and conduct. Departments were also reluctant for more pragmatic reasons. The machinery of government changes of 1987 had disrupted progress with corporate planning: program structures were in a state of flux while, even more importantly, no corporate management process for dealing with evaluation results had been implemented.

Of course, a number of the fears countenanced by line departments were not without foundation and Task Force discussions with supply divisions within the Department of Finance confirmed this. Understandably, many officials in the supply divisions were interested in how program evaluation could be integrated with the budgetary process and, more precisely, the potential of evaluation to assist them in identifying savings. In May 1988, for example, estimates memoranda went to considerable lengths to clarify the position of supply areas:

> Supply Divisions should therefore make clear to departments that provision of adequate information demonstrating department's capacity to implement, monitor and evaluate new proposals will in many cases be an important element in our assessment of the proposal (Estimates Memorandum 1988/14. May 1988. FD: 88/0621).

No doubt this line reflected the imperative of expenditure reduction that marked this time, the burden of which lay heavily on supply areas. However, supply officers doubted their capacity to assist departments in the actual conduct of process and impact evaluations, and believed that their strength lay in advising on program priorities (FD: 88/0621).

The work of the Evaluation Task Force had remarkably solid support from both the bureaucratic and political leadership. At the senior bureaucratic level several pivotal appointments had been made in the mid-1980s that helped to accelerate the pace of administrative change. In early 1986 Michael Codd was appointed Secretary to the Department of the Prime Minister and Cabinet and Michael Keating was assigned the headship of the Department of Finance. Both maintained an unswerving commitment to managerial prescriptions for purported public sector ills (Mills 1992, p25; Zifcak 1994, pp21, 163). Importantly, this extended to a close interest in program evaluation policy. As noted earlier, while Keating was responsible for establishing the Task Force, the authority with which it undertook its task was derived from a joint recognition of its priority (Ministerial Brief, April 1988. FD: 88/0621).

This position was, of course, vigorously shared by successive Ministers for Finance, but perhaps benefited most significantly from the stability provided by the relatively long tenure of one Minister, Senator Peter Walsh. Between 1984 and 1990 Walsh provided both continuity of forceful ministerial style and, in light of his self-confessed preoccupation with fiscal rectitude (Walsh 1995, pp102-105; Kelly 1994, p490), considerable sympathy for the potential of evaluation policy initiatives to support central budgetary processes. Indeed, at the initiation of the Evaluation Task Force, Walsh had written to the Prime Minister in the following terms:

> Consistent with the establishment of mega-departments and portfolios and the associated budgetary approach of allowing Ministers some scope in priority setting through portfolio targets, it is particularly important to enhance the quality of evaluation activities on a systematic footing. Accordingly, I wish to seek support among our Ministerial colleagues for the work of the Task Force, and to give that work a high Government profile (Finance Minister to the Prime Minister, April 1988. FD: 88/0621).

In June 1988 the cabinet noted the objectives of the Evaluation Task Force and reaffirmed the program evaluation planning process introduced in September 1987. In September the Minister for Finance submitted to cabinet a proposal for a government-wide evaluation strategy. In the context of the introduction of portfolio budgeting during the 1988 budget, which assigned portfolios greater responsibility in reviewing programs and meeting ERC determined portfolio expenditure targets, systematic program evaluation was seen as “particularly important” in serving a dual role and providing

> a more rational basis to assist priority setting and resource allocation within portfolios as well as for ERC deliberations (FD: 88/0621).

A critical proposal related to the role of the Department of Finance in enhancing evaluation capacity in portfolios. This required Finance to act as a “catalyst and coordinator” for program evaluation by producing guidelines, making available Finance resources, disseminating evaluation skills through participation in joint evaluation exercises and assistance in developing appropriate management information systems to support evaluation. However, following discussions with departments, it was agreed that this role required further clarification. While existing evaluation planning requirements sought to locate responsibility with portfolios, the central agencies also had legitimate responsibilities for evaluation which derived from their “broader budgetary and cabinet coordination roles”. As a consequence planning requirements were revised to include only those evaluations with major resource or policy implications, or were likely to require cabinet or joint ministerial consideration.[7] In these cases, it was proposed that

> central agencies seek to ensure that portfolio evaluation priorities identified in evaluation plans are consistent with agreed government wide priorities and polices, and also to take portfolio planning into account in the consideration of the budget review agenda (FD: 88/0621).

It was also proposed that central agencies—and in particular, the Department of Finance—be allowed to comment upon terms of reference for major evaluations and participate with portfolios on selected evaluations.

[7] The factors to be taken into account when selecting “major” evaluations included the following: (a) the significance of the program in terms of materiality and political significance; (b) whether the program has been specifically identified by the government for evaluation; (c) known problems in terms of either implementation or program impact; (d) changed circumstances of the program; and (e) a desire for a reasonable distribution of evaluations and their coverage (FD 88/0621/4).

While the planning process remained central to Finance's evaluation initiatives, there were also dangers in relying solely on this mechanism. In reviewing the cabinet submission, division heads identified a number of possible tactics that departments could employ to circumvent underlying objectives and so "frustrate options for budgetary restraint". These included ploys such as placing the "more vulnerable" programs at the end of evaluation plans and tying supply areas to committed evaluation priorities that may not sit well with prevailing budgetary imperatives (Division Heads Meeting, September 1988. FD: 88/0621). The issue facing Finance was how to minimise the incentives for this kind of adaptive behaviour.

Of course perseverance with the evaluation planning approach was not universally accepted by departments and this was duly reflected in their coordination comments on the September cabinet submission. A number were disappointed that the resourcing of evaluation and evaluation planning was dependent on existing and inadequate program funding levels (DASETT, DAS, Defence, DPIE, DSS). By far the most common complaint was with the implicit logic of evaluation planning which did not sit comfortably with other reforms, especially devolution. Finance guidelines were perceived, with some justification, as "unjustified incursion into portfolio responsibilities" (DFAT, DEET, DITC, DTC, Treasury, ATO, DVA) and sought to usurp the portfolio budgeting "prerogatives" of Ministers (DSS, DEET). However, the most incisive observations came from Attorney-General's and the Department of Immigration, Local Government and Ethnic Affairs (DILGEA), both of which noted how evaluation planning had been quarantined from, rather than integrated with, program management and corporate planning. In particular, DILGEA feared that

> [t]he recommendations would lead departments and agencies to develop evaluation plans which have no meaningful relevance to operations (Coordination Comments from Departments on the Cabinet Submission "Enhancing Program Evaluation in the APS". FD: 88/0621).

For these two departments the mere fact that evaluation planning requirements had been effectively excised from corporate management planning showed clearly the transparency of Finance's true intentions for evaluation. Evaluation planning was distinct because it was earmarked for distinct—that is, budgetary—purposes.

The Evaluation Task Force reported in October 1988.[8] Its primary finding was that program evaluation in the APS was deficient across a number of areas (although given the tenor of portfolio comments described earlier as well as the relatively short lead time given to departments in terms of guidelines for the planning of evaluations, this was not particularly startling). The Task Force adopted the line that evaluation was the key to redirecting the focus of budgetary decision making towards "effective outcomes and value for money" because the rationale of evaluation forced decision makers to consider factors other than resources. The most propitious route to securing an evaluation "culture" was to establish a systematic and coordinated evaluation strategy that would, in tandem with program budgeting, encourage departments to determine their priorities: the ultimate objective being to improve services to the public (Evaluation Task Force 1988, pp2-3).

In consultations with departments, and as outlined above, the Evaluation Task Force distilled a number of problems or misconceptions which required rectification. Of most import was the embryonic stage of performance measurement in most departments; wildly varying definitions of what constituted evaluation activity and, as a consequence, a heavy bias in evaluation activity towards process and efficiency rather than effectiveness and "continued relevance"; apprehension about the adequacy of existing evaluation skills and analytical capacity within departments; a sense of obfuscation as to the role of central agencies and, as a corollary, an attendant preoccupation of senior program managers with the cost-cutting or "punitive" aspects of evaluation (Evaluation Task Force 1988, pp4-5; Barrett 1992, p5).

The thinking of the Task Force recognised this by concentrating on three areas. The first was accelerating the development of performance information systems within departments as an information base to assist evaluation exercises. Second, following the lead of the *Handbook*, it sought to reinforce evaluation as a continuum which moved from process and efficiency to questions of effectiveness. In this way evaluation was cast as a decision making tool for managers at all levels, even though different types of evaluation were more appropriate at different levels (for instance, effectiveness evaluation is more

[8] Although, as an audit of the evaluation strategy later noted, the Task Force did not actually produce a report. Concerned as it was with "broad" policy development the Task Force was apparently of the view that it need only outline its proposals in a "few brief paragraphs" (ANAO 1991a, para 4.5.3, p98).

useful than process evaluation in allocating resources between programs). Third, in order to ensure utilisation of evaluation results, particularly in the budget process, the link between program budgeting and evaluation had to be strengthened (Evaluation Task Force 1988, pp7-12).

In response to these perceived gaps the Task Force prepared a comprehensive "evaluation strategy" for the Government, adopting a broad definition of evaluation as including not only strategic and tactical evaluation as outlined by the *Handbook* but also other forms of review such as efficiency audits, parliamentary committees of inquiry and *ad hoc* boards of review (DoF 1989a, p16). This definitional largesse was designed in part to reduce the intimidating aspects of evaluation as a technique by reminding line managers that evaluation, in the common sense at least, has long been part of government. The evaluation strategy comprised four service-wide components:

> (1) the development of agency evaluation plans (AEPs) by departments for the systematic evaluation of all their programs over a five year period;
>
> (2) the preparation of portfolio wide evaluation plans (PEPs) endorsed by the Minister and forwarded to the Department of Finance. The design intention being to "better integrate program evaluation within central budgetary processes";
>
> (3) the strengthening of evaluation reporting requirements; and
>
> (4) the implemention of measures to improve the level of evaluation skill throughout the Australian Public Service (DoF 1989a, pp16-17).

In mid November 1988 cabinet endorsed the evaluation strategy and sanctioned the Department of Finance's coordination and development role (DoF 1989a, p16; Cabinet Minute No.12029. FD: 88/0621). At this time Finance distributed a memorandum to all portfolio secretaries outlining the formal requirements for planning.

As part of the sales pitch of the strategy, program evaluation was seen primarily as the responsibility of individual departments. As we have seen, however, Finance had been charged with a co-ordinating function consistent with its role in both general expenditure control and framework management for the FMIP. It was expected to reconcile the priorities identified in PEPs with government-wide policy priorities, comment on the adequacy of terms of reference for evaluations and participate on steering committees where required (DoF 1989a, p17; Barrett 1988, p56). Despite the dovetailing of evaluation and portfolio budgeting, the Department of Finance was adamant that "the Government requirement that PEPs be produced" meant that evaluation

> is not . . . seen as the exclusive province of individual Ministers and their departments. Managers can not expect to be only accountable to themselves (DoF 1989a, p18).

As a consequence, from the inception of the evaluation strategy the Department of Finance was acutely aware that a

> fundamental challenge for Finance is to ensure compatibility between the role of evaluation within departments and its role in the central budgetary processes: care is needed to ensure that the latter role does not discourage the former (DoF 1989a, p17).

This tension also loomed large in the minds of senior Finance officials who sought to play down the strong budget connections (see, for instance, the views of former deputy secretary Pat Barrett: Barrett 1988, p56). Indeed, it was the relations between supply officers in the General Expenditure Division (GED) of the Department of Finance—who are responsible for examining the adequacy of program expenditure proposals—and other departments undertaking evaluation which exposed the most "potential conflict" (DoF 1990a, p22). This potential was accentuated because supply areas in Finance were given a crucial lead role in monitoring and, in the drafting process, assisting the portfolio evaluation planning process initiated by the November cabinet decision (Budget Circular 1988/136. FD: 88/0621/4).

The shape of this tension—priority setting in the budget context—will be discussed more fully in the next chapter. For the moment it need only be noted that Finance attempted to reduce the potential for conflict via two developments. The first of these was a change in nomenclature. In 1988 program budgeting was changed to the more inclusive "program management and budgeting" so as to reflect the fact that managers were managing *programs* in addition to budgets (DoF 1988a, p12). The second development mirrored the first. In 1989 two additional evaluation guides were prepared by the Evaluation

Task Force of which the most significant was *Program Evaluation: A Guide for Program Managers* (hereafter the *Guide*). The *Guide* represented a discernible shift in the orientation of the evaluation strategy, away from a mechanistic budget driven process. In stark contrast to the earlier *Handbook* the chief purpose of evaluation was to

> help managers to plan and implement their programs effectively and efficiently. This serves to ensure that program objectives and client needs are met in a manner consistent with government objectives (DoF 1989b, Foreword).

This adjustment in evaluation's focus towards the client, combined with a more forthright approach to the use of "qualitative" methods in program evaluation, tend to suggest that the *Guide* was a direct descendant of the government's belated embrace with social justice issues at the tail end of the 1980s (DoF & DPMC 1989; see also DoF 1989c; Boston & Uhr 1996, p61).

However, the *Guide* is perhaps of more significance for the articulation of three categories against which government programs should be evaluated: efficiency (where program inputs are used so as to maximise program outputs); effectiveness (where program outcomes achieve stated objectives); and appropriateness (where program objectives match government priorities and community needs) (DoF 1989b, p1). Of these, appropriateness carries the most weight if only because it is the most daring. By asking program managers to assess the appropriateness of their programs, the *Guide* was conceding that "when properly performed, evaluation calls upon policy and political judgments by administrators" (Uhr 1990b, p89). In this sense, the appropriateness criterion was the most redolent of the RCAGA's agenda of infusing the public service with a more explicit policy role. At the same time, however, this concession may be empty if program managers were rarely called on to evaluate the appropriateness of programs. And the scope of these responsibilities will, in turn, be related to the extent to which the evaluation strategy is orientated towards central budgetary decision making. Here then resurfaces the underlying structural tension in evaluation policy that has been traced from the early design period—evaluation serving fundamentally incongruous purposes. As the next section will show, the implementation phase of the strategy reveals the processes of resolving—or obfuscating—this tension.

Implementing The Evaluation Strategy

Given the indifference, if not at times outright hostility, displayed by policy makers and government departments over the last twenty years to questions of systematic evaluation, it is not surprising that the implementation of the government's evaluation strategy has been portrayed by a number of senior officials in the Department of Finance as essentially a task of effecting "cultural change" within the APS (Sedgwick 1993, p1; Interview 03). From the inception of the strategy—and as evidenced early by the first set of planning guidelines—Finance has consciously adopted a "softly softly" approach to enforcing portfolio compliance with the strategy (Interview 15; cf ANAO 1991a, para 4.6.1, pp104-5; para 4.9.2, pp129-30). Program evaluation was primarily the responsibility of individual portfolios. This was not to deny the bureaucratic reality that compliance was only likely to be secured if evaluation *planning* obligations were cast as mandatory, complete with the sting of penalties, but rather a recognition that compliance with planning requirements was not an end in itself. The chief consideration was that a bureaucratic requirement to submit a report to the Department of Finance compelled portfolios, at the minimum, to establish structures (Interview 15). However, evaluation policy was the *sine qua non* of the FMIP focus on results; securing process alone would ultimately have been self-defeating.

The primary responsibility of the Evaluation Task Force was to advise portfolio evaluation units on evaluation, develop and present training courses, and provide assistance on particular evaluations. However, the Task Force was always intended to be a transitionary instrument for fleshing out an evaluation policy and generally increasing a service-wide awareness of the importance of evaluation to the government's reform program. After the expiration of the Task Force in late 1988, its evaluation function was superseded by the establishment of the Resource Management Improvement (RMI) Branch within the GED (Interview 15). The RMI Branch, as its title suggests, was assigned responsibility for policy development in the broad area of management improvement across the service. This incorporated aspects of performance information, program budgeting, cost-benefit analysis and program evaluation. Of these, the RMI branch was, at this time, mainly occupied with the provision of advice on cost-benefit analysis, while the program evaluation section was concerned with providing *ad hoc* assistance for individual portfolios on particular evaluation exercises. Throughout this period the Principal Adviser in

GED was assigned a coordinating responsibility for the strategy (ANAO 1991a, para 4.4.2, p90). This arrangement was to continue until January 1991 when the evaluation strategy received a permanent institutional home with the creation of the Evaluation and Statistical Services Branch (ESSB), again within the GED. According to several long serving officers this was overdue recognition that “evaluation was a big enough topic to warrant a separate unit” (Interview 20, 23).

The immediate objectives for ESSB were to provide both “a focus for the Department’s evaluation facilitation activities” and “clearer and stronger links with the Department’s Supply Divisions and with other departments and agencies” (DoF 1991a, p86). The dual structure of the RMI branch was maintained, with separate sections having responsibility for economic evaluation methodologies and evaluation policy development. However, the utility of an economic evaluation section, specialising in cost-benefit analysis, had by early 1991 diminished considerably. Most departments had built up competencies in this area and reserve Finance capacities were seen as redundant (Interview 10, 15). As a consequence the economic evaluation section was dissolved in favour of a dual structure focused on the branch’s “facilitation activities”.

The policy development work anticipated for the second section of the branch did not materialise for the period between 1988 and late 1992. Not unexpectedly, implementation of the strategy within portfolios was the first priority and no major, internally initiated policy changes were made (Interview 23). The branch was now fully occupied with the task of promoting program evaluation within portfolios through seminars and the preparation of handbooks and guidelines, providing technical assistance for particular evaluations where requested, assisting portfolios in the planning process, and liaising with portfolio evaluation units. For these purposes the sections of ESSB were organised into two evaluation consultancy groups (ECGs) and each allocated responsibility for “shadowing” a specified number of portfolios. The deliberate use of the term “consultancy” in the title accurately captured the role of ESSB in advancing the evaluation strategy. Portfolios received assistance only when they requested it. The Department of Finance was, in this respect, merely the “sponsor” for program evaluation (Mackay 1994a, pp18-19).

A REALITY CHECK ON IMPLEMENTATION

Of course the methodical development by Finance of implementation structures and the application of guidelines provides only one perspective on the course of the government's evaluation strategy. There are, however, other perspectives that have served constructively as a 'reality check' on implementation. These identify discrepancies between Finance's modelling of an evaluation strategy and aspects of program evaluation as practiced by portfolios. As noted earlier, the Department of Finance deliberately adopted a "softly softly" approach to encouraging program evaluation activity in portfolios in preference to a heavy handed compliance approach. The reasoning was that making portfolio evaluation planning mandatory was incentive enough for portfolios to structure their internal evaluation efforts. However, progress with evaluation *within* portfolios has been monitored almost exclusively by external examiners, particularly the Australian National Audit Office (hereafter Audit Office). The series of efficiency audits completed by the Audit Office in the early 1990s is also crucial because it represents the foremost—if not only—source of publicly available evidence on the implementation phase of the evaluation strategy that is also *external* to the agency responsible for its inception. These circumstances certainly give shape to the Department of Finance's rhetorical commitment to devolution. In truth, however, this inattention is also a function of the underlying budgetary priorities of the evaluation strategy.

Developments in evaluation policy had been noted with interest outside government administration. An instructive early example is the report of the House of Representatives Standing Committee on Finance and Public Administration into the implementation of the FMIP, *Not Dollars Alone* (HRSCFPA 1990). As the title suggests, this inquiry was, in part, a response to criticism, emanating particularly from public administration academics, that the FMIP—or 'new managerialism' as it was pejoratively referred to—was pursuing efficiency at the expense of established accountability and social equity objectives. The opportunity to reassert parliamentary claims to evaluation as an accountability tool was particularly inviting and the committee picked up where predecessors had left off. Although the evidence before the committee indicated that evaluation planning processes were "still at the developmental stage" and would require "consistent and significant effort" in incorporating them into departmental corporate and program management practices, the committee was confident of the potential benefits of evaluation

(HRSCFPA 1990, paras 7.52-7.54, pp84-85). These benefits were most likely to accrue if the effectiveness information from evaluation reports were linked more closely to mechanisms of parliamentary scrutiny (that is, functional committees) (HRSCFPA 1990, paras 8.39-8.41, p96). As a consequence the committee made two crucial recommendations. First, that departments make a "consistent and systematic effort" to develop evaluation planning and ensure that this is linked to decision making. Second, that the Department of Finance

> increase and more effectively target its practical assistance on evaluation to departments; and ensure that departmental evaluations are linked into the decision-making of the central budgetary process (HRSCFPA 1990, para 7.62, pp86-7).

On the first recommendation there was to be "demonstrated improvement by 1992" through an "independent effectiveness review of the FMIP" (HRSCFPA 1990, para 10.38, p126). This was restating the position taken by a Senate committee in an earlier report that past efficiency improvement programs (such as the ESU) had not been adequately reported on publicly (SSCFPA 1989, paras 5.22-5.24, p53). Obviously, as discussed earlier, the second of these recommendations was a spur to the creation of a specialised unit—ESSB—within Finance.

To some extent the committee's concerns over program evaluation policy had been anticipated by the government in mid 1990, as evidenced through two separate cabinet decisions. The first decision, in May, was an attempt to tighten the evaluation-budget link and underpin portfolio budgeting. In preparing evaluation plans, portfolio ministers were required to examine their plan and evaluation terms of reference to ensure that the next three year's evaluations provide a better informed basis on which to assess both possible savings and new policy (Cabinet Minute No.13765. FD: 88/0621). The second decision of July 1990 was also preemptory in establishing guidelines for the circulation and public release of evaluation results and reports, and making it a government requirement that program evaluations included in PEPs should "normally be released publicly" (Cabinet Minute No.14025. FD: 88/0621/4; guidelines are reproduced at DoF 1991b, p7). Publication was advocated on the basis that it

> increases the transparency of programs and helps to promote greater public accountability and to facilitate Parliamentary scrutiny. . . It also fosters greater public interest and more informed comment on

> Government programs, and helps to achieve greater public understanding and acceptance of the reasons for change (DoF 1991b, p1).

Public scrutiny was, in the majority of cases, catered for, but noticeably only for PEP evaluations.

In the meantime monitoring of the implementation of the evaluation strategy had been left to the Audit Office. This responsibility was developed as part of the Audit Office's general efficiency auditing function and stems from recommendations made by the Joint Committee of Public Accounts in a 1989 report on the role of the Auditor-General (JCPA 1989; ANAO 1991a, para 2.4.1, p26). The JCPA report was an acknowledgment of the limitations of a devolved administrative environment in which program managers were assigned responsibility for assessing the effectiveness of their programs. It emphasised that there was "room for neutral agencies" to scrutinise manager's evaluation plans and ensure that they have satisfactory procedures in place to evaluate and report on program effectiveness (JCPA 1989, para 11.88, p156). Accordingly, the report recommended that

> the Auditor-General give priority to development of the capacity to criticise constructively auditee's evaluation plans without commenting on the merits or otherwise of government policies (JCPA 1989, para 11.93, p157).

This recommendation was the consummation of proposals aired during the introduction of efficiency auditing in 1979, and mentioned briefly in the previous chapter. The Audit Office welcomed this mandate and wasted no time in initiating a series of efficiency audits of the evaluation strategy in early 1990. The results of these audits were to have a substantial impact on the course of the implementation process.

In April 1991 the first report of a two stage efficiency audit of the implementation of the evaluation strategy was completed by the Audit Office (the second stage report, concentrating on the outcomes of evaluations, will be examined in the context of the budgetary process in the following chapter). The weighty report is an important marker on the evaluation map because it teases out many of the design tensions that emerged during evaluation's gestation period. The audit's brief was to assess the performance of both central agencies and portfolios against the framework established by the evaluation strategy (ANAO 1991a, para 2.4.4, p27).[9] It approached this task methodically, the somewhat sober conclusion being that

> significant progress had been achieved since 1987 but the extent of this progress could not be considered satisfactory in terms of the objectives set by the Government (ANAO 1991a, para 1.3.1, p8).

The constraints on progress could be found at both central and portfolio level. At the central level, the Department of Finance's "softly softly" facilitation role was the subject of stinging criticism. In the Audit Office's view Finance had refused to "enforce compliance" with its guidelines and had, in fact, permitted gross departures. In effect the evaluation strategy received special treatment since Finance had departed from its usual "enforcer" approach to implementation tasks (ANAO 1991a, para 5.1.13, p136). This had the "unintended effect of slowing the growth of the strategy" and undermined Finance's ability to monitor and measure progress with the strategy (Bowden, Ballard & FitzGerald 1991, p5). Moreover, the Department of Finance, as the agency responsible for both resource allocation and the coordination of the evaluation strategy, should have done more to make its position less ambiguous. The latitude displayed in enforcement policy sent mixed signals to departments and fuelled "confusion" about the priority being given to evaluation by the Government (ANAO 1991a, para 1.2.14, p7; para 5.1.10, p135; para 5.1.13, p136). Failure to clarify this promoted reluctance on the part of departments to seek Finance participation and advice, "suspicious of the added advantage Finance might gain from evaluation" (Bowden, Ballard & FitzGerald 1991, pp5, 9-10).

This unease on the part of departments was reflected in the variegated responses in implementation. Evaluation planning was, in most cases, unfocused, incorporating too wide a range of evaluation-type activity. This was attributed by the Audit Office to the imprecision of the strategy's definition of 'evaluation' which had encouraged sloppiness (Bowden, Ballard & Fitzgerald 1991, p5). At the same time, however, the strategy presided over areas of

[9] Six portfolios were examined that spanned "a range of portfolio types and evaluation aspects": Employment, Education and Training, Finance, Foreign Affairs and Trade, Prime Minister and Cabinet, Transport and Communications, and Veterans' Affairs (ANAO 1991a, para 2.4.7, p27).

neglect, especially in the evaluation of policy advisory and regulatory programs which the Department of Finance conceded had been "consciously deemphasised" in the early phase of the strategy (ANAO 1991a, paras 3.2.8-3.2.13, pp32-34; para 4.3.7, pp85-6; para 4.6.11, p108).

In general, the evaluation process within portfolios was in particular need of refinement. Distilling recommendations from the practice of better portfolios, the audit report emphasised the importance of central management support for evaluation exercises, not only in coordination but in the management of topic selection, steering committee operation, adequate resourcing, monitoring and mechanisms for considering and implementing evaluation recommendations (ANAO 1991a, para 5.1.22, pp138-9).[10] Just as importantly, clear responsibility for evaluation had to be assigned to either a senior officer (preferably a deputy secretary) or a committee of senior management. This, however, would not be sufficient to ensure quality control in evaluations. A separate mechanism was preferred by the Audit Office for assessment by people independent of the evaluation. This was in accordance with standard Audit Office practice (ANAO 1991a, para 5.1.22, p138; Bowden, Ballard & FitzGerald 1991, p8).

As an aid to improving the strategy's implementation the report made two cutting recommendations. First, that the focus of the evaluation strategy should shift appreciably from implementation to quality assurance in, and the utility of, program evaluations. A prerequisite for this was for Finance to "review its administration with a view to achievement of closer adherence service-wide to

[10] The Audit Office produced two further reports that reviewed progress with program evaluation within three previously unexamined portfolios, Social Security (DSS) and Primary Industries and Energy (DPIE) (ANAO 1992a), and Industry, Technology and Regional Development (DITRD) (ANAO 1993). The second of these audits was considerably less deprecatory than the two initial reports, noting that DITRD had attained an "appropriate stage of implementing the strategy" and succeeded in integrating evaluation within program management (ANAO 1993, p5). However, both audits reported similar procedural misgivings to those outlined above. A failure adequately to control evaluation planning and conduct, resulting in unsatisfactory on-time completion rates, as well as poor reporting and evaluation utilisation mechanisms were identified as problems in the large DSS and DPIE portfolios (ANAO 1992a, pp2-3, 20-25, 31-32, 35-39, 45, 57-64). In the case of DITRD, problems included insufficient attention to procedures for monitoring the quality of internal evaluations against pre-determined standards, wide variations in the reporting and presentation of evaluations, inadequate attention to cost control of evaluation activity (although this was widespread) and the absence in several portfolio agencies of mechanisms for ensuring the implementation of accepted evaluation recommendations (ANAO 1993, pp14-20).

basic design features of the framework” (ANAO 1991a, para 1.2.11, p6). In particular, although the PEP process was touted as the primary gauge for compliance with the evaluation guidelines, the report’s analysis of the first three rounds of planning (1987-1989) indicated that portfolio evaluation planning was deficient in a number of fundamental areas. These included lateness in lodging plans, widespread absence of ministerial endorsement of plans, and a failure to report on completed evaluations or indicate priorities for future evaluations (ANAO 1991a, paras 3.5.8-3.5.40, pp61-71). As a consequence, Finance’s coordinating capacity with regard to evaluation was singled out for further development (ANAO 1991a, para 1.2.13, p7; para 4.6.34, pp113-114).

Second, the Department of Finance was charged with a failure to practice what it preached. Unlike the prescriptions contained in the evaluation strategy, the strategy itself was developed and implemented devoid of performance measures to indicate its impact (ANAO 1991a, para 4.7.2, p114). As the audit revealed, analysis of PEP submissions was not only partial in the intelligence it provided, it also appeared somewhat disingenuous since it constituted a process measure. As a consequence, the Department of Finance was chided to devise some way of “objectively” evaluating the effectiveness of the evaluation strategy (ANAO 1991a, para 1.3.5, p9; para 4.7.20, p120). Given the paucity of performance information it should take more responsibility than hitherto for the collection of appropriate data relating to the impact of evaluation.

There is some irony in the complexion of the Audit Office observations. The usual practice of external assessors is to criticise the authoritarian posture adopted by central budget agencies. In the case of the evaluation strategy there is some truth in Finance’s response that there was “a fundamental philosophical difference of opinion” between itself and the auditors (ANAO 1991a, para 4.9.2, pp129-30). Although acknowledging the prevailing managerial climate of devolution, the Audit Office was “mindful that a devolutionary approach to evaluation also carries a commensurate requirement for coordination and strategic control” (ANAO 1991a, para 3.4.75, p58). As a consequence central agencies, and in particular the Department of Finance, have a “duty” to align evaluation activity with these priorities (ANAO 1991a, para 4.6.7, p107).

Finance, however, did not accept a monitoring role and, by extension, responsibility for the “integrity of the evaluation framework” willingly. This uneasiness was a product of competing intentions: a desire to sell evaluation as a crucial part of the devolutionary message of program budgeting and improved program management, and, as we have seen, a fairly consistent association of evaluation with the budgetary process. The audit report was diplomatic in describing this as a problem of "communication" (ANAO 1991a, para 4.4.5, p92). This certainly presented Finance as wavering, but it was probably more symptomatic of the change of orientation being introduced *within* the department, and particularly the role of supply areas, as the audit touched on. If there were no performance measures for the strategy, this was likely to reflect the "lack of congruence" within Finance as to the objectives of the strategy (ANAO 1991a, para 4.4.15, p94; paras 4.4.22-4.4.23, p96).

One effect of this clash of perspectives was that the Department of Finance did move to repair a perceived communication breakdown. It advised the Audit Office in unequivocal terms that it was "taking action to link evaluation more closely to budget decision making" (ANAO 1991a, para 4.4.13, p93; see also para 4.4.5, p92). To facilitate this Finance conceded ground on the issue of framework integrity—and pointedly with a budgetary bent—by indicating that it would "report more fully on progress" of the strategy and that its focus would "increasingly be on the quality of evaluations undertaken and the use made of results in the budget context" (ANAO 1991a, para 3.4.32, p47). This response foreshadowed a reconsideration of the evaluation strategy in late 1992.

REORIENTATING EVALUATION POLICY

The period of evaluation policy reappraisal coincided with a rapidly changing constellation of political, economic and bureaucratic elements. As a point of departure, in December 1991 political stewardship changed as Paul Keating succeeded Bob Hawke as leader of the Labor Party and Prime Minister. This heralded not only a shift in political style—with Keating appreciably more "aloof from his ministers and detached from his tasks" than his predecessor (Kelly 1994, p xviii)—but also a reversal in policy orientation. Disenchanted by the apparent futility of significant budget deficit reductions on the intractable balance of payments problem, Keating initiated a new policy posture that sought to reinstall social justice as the sole preserve of the ALP and mitigate the problems of high levels of unemployment and stultifying recession. The policy instrument was expansionary fiscal policy which would produce a series of increasing budget deficits during the early 1990s (Kelly 1994, pp xii-xiii, 490-

492). This shift, combined with a comprehensive social and cultural agenda, dissipated the strength of previously sacrosanct economic imperatives.
With the ascension of Paul Keating as Prime Minster also came a cross-over in the bureaucratic leadership. In early 1992 Michael Keating moved from Finance to head the Department of the Prime Minister and Cabinet. His successor was Stephen Sedgwick, an officer imbued with a similar appreciation for the virtues of strong managerial structures, but tinted perhaps with a more pragmatic outlook. The influence of this change on the direction of evaluation policy should not be underestimated. While both are economists and hence share an awareness of the demands of the finance ministry, Keating displayed less confidence in the equation of budgetary with political or policy considerations, especially at cabinet level, while Sedgwick was somewhat more receptive to the purported benefits of this type of budgetary information (Keating 1995, p3; Sedgwick 1992, pp45-7). Career paths may partially account for the difference: Keating was resolutely a career public servant while Sedgwick had moved from senior ministerial advising positions into the public service, suggesting that Keating may have held more subdued expectations for what was administratively feasible (Mills 1992, pp25, 28-29). Hints of this difference can be discerned in the post-audit reorientation of evaluation policy. Given the sea-change in economic and policy imperatives, the *categorical* statement by Finance of its position on evaluation and the budget can be seen as something of a disjuncture in the developmental process.

The critique of the strategy offered by the Audit Office was not an isolated example. Similar concerns were raised in the systemic evaluation of the FMIP undertaken in 1992 by the MAB Task Force on Management Improvement.[11] Although the countenance of this review was slightly more self-effacing than the efficiency audit it was, at the same time, also somewhat self-delusory, for while it could confidently boast that there was "widespread acceptance of the importance of evaluation" among portfolios much of its analysis tended to suggest otherwise (MAB-MIAC 1993a, pp377-378). Most prominently, equivocal staff surveys cast doubt on the achievements of evaluation as 'cultural change'. In replicating attitudinal research undertaken for the 1984

[11] Established in April 1992, the evaluation was as an initiative of the MAB that sought to overview the reforms, evaluate their outcomes and suggest future directions. It sought also to answer the HRSCFPA's call for "an independent effectiveness review of the FMIP" (MAB-MIAC 1993a, pp1-2). Formal independence, however, was eschewed in place of a review framed as part of ordinary program evaluation obligations (Rogers 1993, p371).

FMIP Diagnostic Survey, the evaluation found a very slight improvement in senior staff views on the place of evaluation in the work of program management—actually running programs remained overwhelmingly superordinate to evaluating outcomes (MAB-MIAC 1993a, pp367-369). This was corroborated by further survey analysis that showed that the vast majority of officers made “little use” of information derived from program evaluations in the course of their work (MAB-MIAC 1993a, pp372-3). (In fact less than ten per cent of officers claimed they used evaluation “frequently” (p373). But of course utilisation is also a function of the intended user; if most PEP evaluations were being used by the obviously fewer number of officers involved with central budgetary negotiations then this data *may* represent effective utilisation.)

Attitudinal stasis was also evident in the two most common reservations about the evaluation strategy: the availability of evaluation skills and the evaluation of policy advice. On the first, despite the extensive training program initiated by the Department of Finance, the perception of evaluation techniques as “special” continued to linger (MAB-MIAC 1993a, pp375, 378; MAB-MIAC 1993b, pp126, 498, 556). Obviously this was presenting complications for the integration of evaluation as a “normal” part of everyday operations for program managers, even though the review stressed that both attitudinal and structural change was best promoted by resisting the temptation to centralise evaluation skills (MAB-MIAC 1993a, pp377-378). For some social welfare portfolios the technocratic tenor of the evaluation strategy did not provide much assistance:

> Much of the rhetoric and pseudo-science that has arisen around evaluation makes it inaccessible to managers and often produces dubious ‘scientifically accurate’ results. Evaluation must be seen and promoted as a common sense management tool—get managers to see this and formal requests will not be seen as onerous (DHHCS cited in MAB-MIAC 1993b, p272).

Perhaps most distressingly the evaluation could discern signs of a tendency on the part of portfolios to engage in evaluation planning as merely the satisfaction of “bureaucratic requirements” (MAB-MIAC 1993a, p379).

The second area of evaluating policy advice represented something of a leap of faith for the strategy. A major impediment to evaluation in some portfolios—and in particular central coordinating agencies—was the apparent difficulty of applying the strategy to policy advising 'programs'. The Department of Finance was unduly circumspect in describing this component of the evaluation strategy as "an area of great complexity" (MAB-MIAC 1993a, p376). Nonetheless, as Chapter Seven will detail, the evaluation of policy advice was, after late 1991, one of the principal factors in reorientating the focus of the evaluation strategy.

The external prodding provided by the audit findings provided Finance's ESSB with some ratchet policy development work and precipitated a marked change in the emphasis of the strategy (Interview 13). An obvious manifestation of this was a renewed focus on training and skills dissemination evidenced by the initiation of a new *Evaluation Papers* series specifically directed at program managers and evaluators (e.g. DoF 1992a). During late 1992 and early 1993 an intensive reconsideration of the strategy was undertaken (Interview 10, 15). Finance believed it had built up good relations with portfolios, especially on the issue of evaluation planning, and the review process was carefully managed to ensure that this goodwill would not be "squandered" (Finance Minute July 1992: FD: 92/2352). An ESSB assessment concluded that with limited resources ESSB had done a commendable job in promoting evaluation throughout the service: it had taken evaluation to departments and had achieved a "positive outcome". As part of this introspection, a cabinet submission was prepared that suggested that the way ahead was to "let go" but at the same time ensure that the compliance orientation shifted appreciably towards demonstration of the utility of evaluation through the work of supply areas. Even though the first proposal flew in the face of the deafening Audit Office advice, this approach received support from the Secretary, Stephen Sedgwick (Interview 10, 20).

The schizophrenic-like projection of Finance priorities for evaluation—as crucial to both program management and budgetary decision making—that was built into the strategy came to a head. Almost immediately the focus on PEPs intensified. This occurred despite both their overt process orientation *and* the fact that a number of internal reviews of progress indicated that, while compliance had certainly increased, the quality of the plans as strategic documents remained questionable (FD: 91/3905; Interview 19, 23). With it the dual priorities were emasculated so as to concentrate efforts on the use of evaluation in budgetary decision making (Interview 23). Evaluation in the budget was now undeniably the "main game" (Interview 19).

A ramification of policy reconsideration of this type was that the role of ESSB (or the Evaluation and Staffing Analysis Branch (ESAB) as it was renamed in 1993) was also recast. Rather than liaising almost exclusively with portfolios, the function of ECGs was now principally to assist supply areas in the Department of Finance to encourage the use of evaluation by portfolios in budget negotiations and, obviously, stress the importance of evaluation information to the work of supply officers in aligning evaluation priorities and assessing the soundness of policy proposals. Within the confines of the evaluation strategy supply areas were variously referred to as the "front end of the department" (Interview 22) or the primary "gearing ratios" (Interview 15). ESAB assistance came in the form of "buddying liaison". Each officer in the two ECGs was assigned a supply area to shadow. In this way there was an evaluation buddy for each of the portfolios covered by a supply area. This created a direct line of contact between portfolio evaluation units, the responsible supply area and the evaluation branch (Interview 23).

This liaison function was an addition to the existing role of the branch. But the period following policy reappraisal was also marked by a degree of policy disorientation. Ironically this had most application to the role of ESAB. Given its initial position as a sponsor for program evaluation throughout the service, a continuing promotion role for evaluation as a node for information exchange and a training facilitator was being increasingly questioned by some officers as "not appropriate" for a branch that was intended to retain some policy development responsibilities. The training function encouraged a "dependency relationship" with portfolios because Finance's insistence on retaining a training capacity provided a disincentive for portfolios to develop these for themselves. A hint of disillusionment was also being fuelled by an inappropriate staffing structure of relatively senior officers that ensured ESAB resembled a "high powered think tank that is not allowed to think" (Interview 19). However, the catalytic role of the branch implied that, to some extent, it must also have a definite organisational 'shelf life'.

CONCLUSION: A SKEWED POLICY?

One ESAB officer offered the following curt assessment of the evaluation strategy—"the strategy was a series of short-sighted steps which have proved contradictory" (Interview 19). The analysis contained in this chapter indicates that there is some accuracy in this judgment. Certainly the economic difficulties of the mid-1980s catalysed the development of evaluation policy and program budgeting, but one can easily discern a number of the design trends traced from 1976, the confluence of which has contributed to the shape of the current evaluation strategy. More to the point, this chapter has demonstrated that conventional critiques of the current framework, as either unduly oriented towards budgetary decision making or blunted by an “internal managerial” form that limited evaluation to executive decision making, tend to reflect an unfamiliarity with conditioning design intentions (Caulley 1993, pp2, 6-7; Ryan 1992, pp69-74). Evaluation as a tool of central direction was always the strongest current. However, even though a dominant function for evaluation had emerged by the early 1990s, because it also purported to serve a *range* of purposes there was an inherent risk that evaluation might fail to accomplish any of them.

The evaluation strategy articulated three broad objectives: first, to improve the information base for program management; second, to assist government decision making and prioritisation, “particularly in the budget process”; and third, to provide evidence of program managers’ stewardship of resources and hence improve accountability to parliament and the public (DoF 1992c, pp6-7; 1994a, p2; Mackay 1993, pp9-10). These objectives capture most of the design trends over the last twenty years. However, the crucial issue to emerge from discussion in this chapter is the extent to which these objectives have equal standing. Policy reappraisal in the early 1990s has extracted a more explicit stance on the part of the Department of Finance which was itself long preoccupied with the “fundamental challenge” of balancing its resource allocation obligations and its sponsorship role for evaluation.

This ‘skewing’ of the expectations for program evaluation has, particularly in the eyes of line departments, much maligned the facilitative role of Finance. This tension has been characterised elsewhere as the problem of Finance “speaking with forked tongue” (Campbell & Halligan 1992a, pp144-57). Further, the Secretary of the Department of Finance has conceded that there is indeed a “natural element of tension” within the reforms and that the annual nature of the budget can send “confusing signals” to program managers about the relationship between budget priorities and the promotion of improved program management (Sedgwick 1992, p44). The transmission of confusing signals is perhaps best illustrated in the evaluation strategy through the requirement for agencies and portfolios to prepare portfolio evaluation plans (PEPs) and agency evaluation plans (AEPs).

From earlier discussion we saw that the government's evaluation strategy is framed around the major evaluations conducted under PEPs that feed into the budget process (Barrett 1992, p17; Sedgwick 1992, p46). PEP evaluations are "intended to focus mainly on outcomes and effectiveness issues" while AEP evaluations focus principally on process or efficiency issues (Mackay 1993, p10; Barrett 1993, pp23-4). However, the overwhelming proportion of current, completed or planned evaluations are contained in AEPs which are "essentially internal working documents", the quality of which the Department of Finance knows very little (Amies 1994, p35). There exists only a single major internal review of evaluation (within the large employment, education and training portfolio) and this indicates that the strategy is achieving equivocal results. While program evaluation was perceived by the portfolio as primarily a program management tool, there was also considerable evidence that evaluation had failed adequately to inform departmental budgetary and policy proposals destined for cabinet (DEET 1994, pp5-6, 23-25). As a consequence, the operational reality of the evaluation strategy suggests that there is a significant deviation between Finance's intentions for the strategy and portfolio practice. As we have seen, this is also corroborated by the staff survey results contained in the MAB-MIAC evaluation of the FMIP, as well as early observations by the Audit Office (MAB-MIAC 1993a, pp372-73; ANAO 1991a, para 5.1.33, pp140-1). The consequences of this deviation for the strategy will be examined in the context of the budgetary process in Chapter Six.

What needs to be emphasised here is that a critical focus on the central budgetary and policy coordination objectives of the evaluation strategy should not be framed in a pejorative way. Information is the lifeblood of government. It is the lot of central agencies to collect and use it for the purposes of their general policy management functions. From its Australian inception, evaluation has been framed primarily as an instrument for central government decision making. Competing purposes have evolved and have conditioned this primary purpose. But, by and large, these remain of second order importance. Here we can note the limitations of evaluation for external accountability, a common criticism being that the evaluation strategy is unsuitable as a mechanism for parliamentary scrutiny of the executive (Hamburger 1992, pp77, 84-5, 89-90; Ryan 1992, p75; cf DEET 1994, pp5, 11). But this is precisely the intended outcome of the dominant design trend traced in this and the previous chapter. External review was a competing function, but the chief purpose of evaluation was to inform cabinet level decision making. To make the strategy stick with departments, evaluation was equated with the budgetary process—talk money and people would listen. The adverse economic conditions that had prevailed since the mid 1980s were propitious for this alignment and, significantly, the objective was pursued with the imprimatur of cabinet. Reflecting on these developments it was apposite that one former Finance deputy secretary rationalised the strategy thus: “it had to be done and we were the best place to do it” (Interview 03).

Chapter Six

Running To Stand Still? Evaluation And Resource Coordination In The Department Of Finance

To this point the thesis has concentrated on the institutional jostling which characterised the development of a program evaluation policy in Australian government. The outcome, as we have seen, was an evaluation system geared towards the information requirements of central government. The task now is to examine the degree of success which the policy attained in serving this purpose. This chapter is the first of two case studies framed around the policy management functions of central agencies, introduced in Chapter Three. From the adduced evidence, both internal and external to the Department of Finance, it will demonstrate that the evaluation strategy has had only a marginal impact on the resource coordination functions of the central budget agency. By documenting changes in the approaches of budget officials, the chapter will assess the impact of program evaluation on the core function of the Finance Department—the provision of policy advice relating to the expenditure priorities of the government.

The chapter proceeds in four steps. First, it will compare the evolving resource coordination role of the Department of Finance with the expenditure control responsibilities of the old Treasury. Has the change been of the magnitude often claimed by public service reformers? Second, in doing so, it will examine the operating environment of the key budgetary policy advisers, the so-called ‘supply’ divisions whose task it is to analyse the appropriateness of any current and proposed expenditure. To what extent have operating environments there altered in response to the claimed increase in the availability of evaluation information? Third, the chapter will sift the available evidence on the influence of the evaluation strategy on budgetary processes. The fourth and final section will summarise the findings and outline their implications for the resource coordination function of the Department of Finance.

REALISING VALUE FOR MONEY IN GOVERNMENT: FROM TREASURY TO FINANCE

The expenditure budget process, according to the doyen of public budgeting analysis Aaron Wildavsky, "lies at the heart of the political process" (Wildavsky 1984, p5). It is through the annual budget that a government realises its policy priorities and determines who will get what resources, the broad purposes for which they will use them and when. The role of the central budget agency is perhaps the most crucial part of this process, in terms of both resource and policy coordination. At the time of budget formulation, these coordination functions are critical because they provide the primary source of advice on the economic, financial, social and political implications of possible policy options.

At its base cabinet is nothing more than a committee of individuals, with limited time and cognitive resources. Since cabinet (or a designated sub-committee, such as the expenditure review committee) is the "committee of final appeal" during budget deliberations, the central budget agency is part of that apparatus designed to minimise the decision making workload of ministers (Heclo & Wildavsky 1974, pp57, 88; Weller & Cutt 1976, pp98-99; Sedgwick 1994, pp2-4). These resource constraints are all the more likely to be stretched during the sometimes chaotic processes associated with budgetary negotiation and decision making. Under these conditions the role of the central budget agency is undisputed—its task is to coordinate the policy questions requiring cabinet consideration via a finely honed budgetary "filtration" process. Of course, individual portfiolios must first set their own policy priorities, but inevitably the areas of dispute both between portfolio ministers and between departments will be considerable. The task facing the central budget agency is to settle most of these disputes *before* they reach cabinet, by deciding which disagreements can be settled at the bureaucratic level through bilateral discussions between departments, and which disputes are more intractable and hence require cabinet attention. It follows that this filtration process calls for a high level of policy judgment on the part of budget officials.

At this level the line between policy and administrative considerations is an especially thin one. The officials who advise on budgetary matters occupy an essential position as the "policy valve" for cabinet, wielding considerable discretion and potential policy making power through their professional judgment. This capacity is most commonly realised through the processes they maintain for the assessment of departmental submissions for new or disagreed policy—the procedures of the budget process are significant because they will help determine the complexion of the budget outcome (Weller & Cutt 1976, p2; Howard 1986, pp62-64). The discretion of budget officials applies to the selection of programs that should be evaluated, the choice of information to be included in briefing papers reviewing expenditure as well as deciding which policy options will eventually receive cabinet consideration. The Labor Government's evaluation strategy is a prime example of how budgetary processes have been altered in order to influence how decisions of this type are made.

The role of the Department of Finance is captured in its corporate mission:

> Serving Australia by promoting value for money in the management of the Commonwealth Public Sector through advice and service to clients (Commonwealth of Australia 1993a, p8).

In this the activities of the Department cover three broad areas (DoF 1992b, p57; Commonwealth of Australia 1993a, pp8-9). The first two relate to the general responsibilities of the Finance Department for public sector financial management. These include administering and advising on financial legislation relevant to government management, such as the Audit Act 1901, and providing financial automated data processing (ADP) accounting services to departments and agencies. The third relates to supporting the executive government through the analysis and evaluation of resource use in order to achieve the optimal allocation of public resources in giving effect to its policy priorities. This is the resource coordination function that provides estimates of the financial consequences of existing and proposed policies within the budget context.

From the corporate mission the key characteristic of resource coordination is the budgetary criterion of "value for money". In everyday usage, value for money refers to something well worth the money spent. In the budget context, although it retains the core of this common sense meaning, value for money is a term derived from the professional literature of accountants and managers, and is the embodiment of public sector management reform under the Labor Government. Value for money refers specifically to ensuring that the results or objectives of government activities are achieved, or are likely to be achieved, with the most efficient use of resources (Parker 1990, pp292-293). It is the companion of program budgeting in that it involves the *explicit* consideration and comparison of feasible policy benefits (outcomes) given the limited resources (inputs) available (cf HM Treasury 1988, p34). To this extent it must be seen as an exercise in determining opportunity costs.

Except in one respect, the Department of Finance maintains that its role as the chief adviser on resource allocation issues has not altered significantly since the split from Treasury in 1976. The exception is the “manner” in which its tasks are carried out such that value for money now provides the standard test for any budget proposal (DoF 1990b, p15, 1991c, p1; CPD, Senate, Estimates Committee A, 25 September 1990, A313-315). As a consequence of changes to the aggregate budgetary control framework—in particular the forward estimates system and portfolio budgeting, discussed in earlier chapters—there has not occurred any categorical expansion in the reach of the Department of Finance, but rather a shift in the *type* of information that Cabinet has asked of the department. Advice, as Finance claims, is framed around the question of relative policy priorities and the effectiveness and continuing relevance of government programs. This exercise in program comparison is aided by the information produced by program evaluation. Although, according to the Secretary to the Department of Finance,

> [t]he trick in ensuring effective Cabinet government is to be explicit about the differences which arise from the exercise in judgement, those which arise from different conclusions drawn from the available evidence, and those which are really disputes about facts. (Sedgwick 1994, p4).

This concedes that a commonly recognised information base, in the form of program evaluation, remains only a starting point. In advising Cabinet, Finance is charged with making evaluative judgments (using evaluation in the common sense) on relative program priorities which can only be made in light of *all* government activities, a view that specific line departments do not have. This coordinating task is the source of Finance's responsibility for providing advice that offers an alternative to that emanating from functional departments—the "devil's advocate" or "second opinion" role.

Assertions that the resource coordination function of the central budget agency has not changed significantly are not without foundation. The evaluative task discussed above resonates with more traditional Treasury considerations (Weller & Cutt 1976; Weller 1977b; Hawker, Smith & Weller 1979). The old Treasury was customarily seen to be the "most influential of Australian federal departments" with its power based on a dual capacity to provide advice on both public expenditure *and* macroeconomic management. In respect of public expenditure, the Treasury fulfilled two primary functions. The first was framed around monitoring expenditure—ensuring that all spending had the correct authority, was commensurate with the amounts appropriated and directed at the intended purpose. The second function was to coordinate government policy and guard against program failure. These translated into the following:

> Treasury officers therefore see their role as a combination of financial estimating and policy evaluation, and they do not explain how they balance the two functions (Weller & Cutt 1976, p39).

In aspiring to both financial and policy evaluation work the Treasury had, by the mid 1970s, "filled the role of policy coordinator by default" (Weller & Cutt 1976, p41; Hawker, Smith & Weller 1979, p132). By virtue of their information gathering capacities at the centre of government (and with little competition from officials in other central agencies) Treasury officials often provided budgetary advice that incorporated commentary on the effectiveness of programs based on purely "subjective and general assessments". These judgments were resented by officers in line departments who argued that Treasury pretensions to "evaluation" were still framed around the traditional financial aspects of programs and neglected the importance of their non-economic objectives. For them, this kind of amateur policy evaluation was beyond the proper sphere of Treasury operations (Weller & Cutt 1976, pp40, 49-50).

Evaluation then has always been a component of central budget agency responsibilities. The present evaluation strategy is not as new a development as first supposed—the change has occurred mainly in the techniques that are now stipulated. Indeed, the change that the evaluation strategy represents could be explained as a part of the defence mechanism deployed by the Department of Finance during the administrative restructuring of the 1980s. This interpretation also gains support from the fact that during the late 1970s and early 1980s the "departmental line" of the Finance Department was often openly dismissive of the place of evaluation in advice on expenditure policy (see Submission from the Department of Finance to the SSCSW, pp2490-2491. Australian Archives Series Number AA 1982/568). Program evaluation could easily be seen as a rationale to *formalise and hence legitimise* the kind of "policy evaluation" role that the old Treasury assumed but could not delineate. In this respect, value for money and program evaluation provides the justification for incursion into the substantive policy jurisdictions of line departments at a time of fiscal stress. It also secures the position of the Department of Finance during a period of significant administrative restructuring.

The Department of Finance has maintained that its fundamental role in the budget process has not altered appreciably since the old Treasury was divided—the portfolio budgeting system permits departments to work within their total budget, while budgetary negotiations between Finance and departments focus on the effectiveness of what a department is doing rather than the "nitty gritty" of expenditure. Outside the administration, however, those involved in parliamentary scrutiny, who might have been quite comfortable with this role maintenance, doubted the capacity of the newer system to deliver effective expenditure control. A prominent example are the Senate estimates hearings which vented dissatisfaction with the orientation of the forward estimates based budgetary system, contending that using the projections as a base for deliberations and leaving the identification of cuts to *individual* departments diminished the likelihood of terminating ineffective policies. In this context evaluation took on a heightened significance, as the following exchange illustrates:

> **Senator O'Chee**—It just seems to me a little bit difficult to ascertain exactly how well the resources are being used, because the importance placed on various policies waxes and wanes. If a policy is being phased out or is not being used as much as was originally envisaged, there seems to be the possibility of some fat lying there in the departmental budget.

> **Mr Eric Thorn—**That is true and, as part of the process of getting to that fat, we are looking to a better integration of evaluations with both the running of the organisation and the budgetary procedure each year. We in Finance will try to nominate evaluations that we believe should be done so that they can be considered by the Government at the appropriate time (CPD, Senate, Estimates Committee A, 25 September 1990, ppA314-A315).

Within this context, program evaluation was a part of the Department of Finance's "standing brief to suggest to the government areas that might be overresourced". This remit provides for what are essentially two complementary processes. The first is informal and is tied closely to the analytical abilities of individual finance supply officers:

> **Mr Thorn—**In the first place, you are relying a lot on the experience of finance officers simply to pick up what is happening out there or to have an idea that something may be able to be done better. It all starts just from a finance officer being close to the subject matter and thinking around it (Ibid, pA315).

These individual analytical capacities represent the core of resource coordination. A number of studies of budget agency expenditure control have confirmed that informal methods of assessment dominate the approach of finance officers in vetting expenditure and policy proposals, and that this is duly reflected in the significant variability of review procedures across different policy divisions (Heclo & Wildavsky 1974, pp14-26, 45-55, 70-71; Weller & Cutt 1976, pp32-35, 43). The second is the formal processes of review associated with program evaluation:

> **Mr Thorn—**The process I have just described, the [officer judged] savings process as distinct from the evaluation process, essentially was the old ad hoc process. The evaluation process that we are now trying to formalise and weave in is an attempt to be more methodical (Ibid, pA316).

The evaluation strategy, as we have seen, was intended to systematise the information flow for supply officials. However, to feed into consideration of the budget, the second process of planned program evaluation remains heavily dependent on the accuracy and judgment of the first informal mechanism—if there is, according to Thorn, "enough lead time, one will work into the evaluation cycle that premonition that there are savings to be got into an area" (Ibid, pA316). The question now is, just how do supply officers view program evaluation in their work?

On The Information Trail: Supply Areas And Program Evaluation

The issue of supply area orientation within the budgetary process is one that has been aired consistently during both the implementation of the FMIP and the development of the evaluation strategy. Several recent assessments have disclosed that the Department of Finance's commitment to the relaxation of detailed financial control in the mid 1980s came only after protracted internal debate (Campbell & Halligan 1992, pp149-151, 155-157; Zifcak 1994, pp117-119, 165). There were, according to one Finance deputy secretary reported by Zifcak, "genuinely two camps here"—those in the resource management branches responsible for steering financial management reforms who were strongly supportive of enhanced delegation to departments, and the so-called "controllers" who reportedly dominated the various supply divisions. For some officers this represented a marked resistance to relinquishing accepted instruments of control or, more specifically, averting the stripping of their *raison d'etre.* For others this was a function of uncertainty on the part of supply divisions as to the scope of their role in a devolved management environment.

Internal fractures widened, as these studies report, because supply divisions were judged to be both less able and less willing to devolve financial authority. There was a lingering perception, especially among officials in line departments, that supply divisions were adhering to superseded modes of control, and preoccupied with costs and details, although on the latter there was a division. Some agencies argued that supply areas remained "unreconstructed 'bean counters'" fixated with administrative detail rather than the effectiveness of program expenditures, while others took the polar position and claimed that Finance's concerns had extended too far into departmental policy prerogatives (Campbell & Halligan 1992, pp149-151). Such concerns about policy incursion are, however, not particularly new (see Weller & Cutt 1976, pp39-40, 49-50). Nonetheless, these studies have demonstrated that there has been considerable misunderstanding, if not outright resistance, on the part of supply divisions in accepting the devolutionary implications of program budgeting and the relationship between the resourcing of programs and their results. Further confirmation can be found in the Auditor-General's first efficiency audit of program evaluation which noted that a "very mixed range of views" prevailed in the various supply areas as to the value of the strategy and that several seemed to dismiss the planning process as "peripheral to their work" (ANAO 1991a, para 4.4.15, p94).

The studies referred to above represent an anecdotal snapshot of supply division practices at the end of the 1980s, at the very outset of the government's evaluation strategy. The purpose of this section is to examine the resource allocation function of the Department of Finance in the light of six further years of experience by updating this picture of supply work. Are these views still valid? How have supply areas responded to the evaluation strategy? In what ways has it influenced their work? As a point of departure let us first describe the nature of supply work in Finance.

The Nature Of Supply Work

The intention here is to confine discussion of resource coordination to those functions performed as part of the department's role in advising the government on how to make the most effective use of its limited budget funding. The Department of Finance itself refers to this task as "budget development and management" and this corresponds with program 1 of the department's program structure (see diagram, Appendix II). The supply divisions, which in total consist of approximately 250 officers, assume sole responsibility for overseeing expenditure across all budget dependent agencies of the Commonwealth government. For this purpose they are divided into four divisions that correspond with (or "shadow") a policy sector within the government—defence and industry, labour and international, social security and transport and government (Appendix II). The four supply divisions are responsible for sub-program 1.2 "oversight and evaluation". In this they have four primary functions:

(a) carriage of estimates processes for the official forecasts of government outlays and revenue, comprising management of the forward estimates,

estimates for new policy and savings proposals, and program and running costs estimates;
(b) analysis of the efficiency and effectiveness of existing major government programs and assessment of policy options for both increased or reduced outlays;
(c) analysis of portfolio ministerial proposals for new policy proposals or savings proposals;
(d) general monitoring of expenditure on approved activities (DoF 1992d, p27; Commonwealth of Australia 1993a, pp30-31).

The supply divisions carry out these tasks in conjunction with the general expenditure division, which coordinates overall budgetary strategies (through its expenditure policy branch) and advises on policy development in the area of public sector financial management (including the evaluation strategy) (DoF 1993a, p11).

Supply work is framed around two nodes. The first encompasses ongoing dialogue with portfolio budget officers which is directed at monitoring ongoing expenditure and maintaining the forward estimates as an information base for Cabinet decision making. For the most part this type of consultation is not formalised. The second refers to the crucial advisory role that accompanies the hectic budget rounds each year. Here, supply divisions are required to analyse the cost-effectiveness of both ongoing programs and new portfolio ministerial policy proposals and advise both the Minister for Finance and Cabinet on their “merits” within the government’s stated policy and budgetary priorities. As the previous chapter discussed, from a budgetary perspective it was precisely because of this continuing contact with portfolios, as well as their appetite for review information, that the supply divisions were charged with the responsibility of promoting portfolio evaluation planning (see Barrett 1988, pp55-56).

Because of the changes in the forward estimates system, the focus here will be narrowed to policy advising on new policy proposals (NPPs) and savings options (SOs) during formulation of the budget. We can note that the nature of supply work is conditioned by a number of structural characteristics that will, to some degree, impinge on the receptivity of supply practices to the output of program evaluation. Of the four functions listed earlier, the last three constitute the policy advising role of supply areas proper. However, estimates and policy advising tasks are inter-related—a large proportion of advising work is directed at estimates preparation—and it is both difficult and hazardous to separate the functions since they all compete for supply area resources. For instance, when advising on new program expenditures, supply divisions are required to provide an independent analysis of portfolio costings (an estimates task which must be settled by both Finance and departments prior to Cabinet consideration) *and* provide a Ministerial briefing assessing the merits of the proposal. Although the processes can be distinguished, according to Finance, these tasks taken together oblige supply officers to

> make judgements on the boundary between costing issues which should be resolved by officials, and policy issues that should be resolved by officials, and policy issues that should be taken up in coordination comments and briefings for ultimate resolution by Cabinet (DoF 1992d, p153).

This is a good indicator of the important “policy valve” role played by Finance supply divisions when negotiating with departments on budgetary allocations. At the same time, however, in those cases where proposals are defined solely by an intention to spend a specified amount of money “the costing itself represents the policy” and the proposing department is likely to claim that the proposal is not open to negotiation (DoF 1992d, pp154-155). Under both sets of circumstances the task for supply officers is to join the “information trail” in order to marshal as much evidence (or performance information) either relating to existing programs or sourced from the general community and affected interest groups as to the need for a program. This is where program evaluation, as one source of information about the effectiveness and appropriateness of government programs, enters the vision of supply areas.

The integration of information derived from program evaluation into the negotiations for budgetary allocations was, according to one deputy secretary, predicated on the capacity of supply divisions to

> indicate priorities from a budget perspective as well as the programs it sees as most “in need” of evaluation to improve resource use efficiency and effectiveness . . . Finance officers would participate directly in certain evaluations to ensure that any subsequent debate

> about resource allocation in the budget context takes place on a more informed and factually agreed basis (Barrett 1988, p56).

Hence there are three ways in which the work of supply areas is likely to have been influenced by the evaluation strategy—the use of evaluations in the policy advice produced by supply divisions, the capacity of supply areas to influence portfolio evaluation planning priorities to fit with the government's (or Finance's) budgetary priorities, and the extent to which portfolios have used program evaluation to place budgetary negotiations "on a more informed and factually agreed basis". Each of these issues will now be examined in turn.

Using Program Evaluation In Supply Work

There is some evidence that the one characteristic common to the work of supply divisions is the degree of inconsistency in applying budgetary procedures. This disparity is a function of the way in which budgetary negotiation is framed around the personal relations between individual supply officers and counterparts in the departments. Earlier studies of supply division methods for analysing policy in the old Treasury described "informal methods of procedure" whereby the budget proposals emanating from different departments were treated quite differently according to the "track record" or reliability of the department in submitting rigorously argued proposals (Weller & Cutt 1976, pp34-35, 43-44). If a department was trusted by the supply area then policy proposals were often subject to considerably less examination than if they came from a department with a more suspect reputation. Individual judgment as to the merits of a proposal inevitably played a dominant role. A similar level of variance was found more recently in the estimates practices of supply divisions of Finance, particularly in the application and interpretation of running costs rules which, consistent with a devolved running costs system, were often based on assessments of prevailing circumstances rather than the uniformity of official guidelines (DoF 1992d, pp95-101).

Given that part of the rationale for a policy of systematic program evaluation was to enable these kinds of budgetary negotiations to take place on a surer footing, by providing commonly accepted information on the effectiveness of programs, there appeared to be significant variation in the approach and attitude of supply divisions to the value of program evaluation to their work. This view was first encountered outside supply areas and most fervently within the Evaluation and Statistical Analysis Branch (ESAB) which has responsibility for program evaluation policy development. Here there was broad agreement that procedural change had been painfully slow. Despite the fact that they have standing relations with departmental program areas—which *should* spill-over into advising on portfolio evaluation priorities and participating on evaluation steering committees—supply areas were "still concentrating on the numbers" rather than using information on program results to routinely assess budget proposals (Interview 19, 23). This contention was reinforced by more senior officers who lamented that supply areas have "taken on evaluation with varying degrees of rigour; unfortunately an outcomes approach is the exception rather than the rule" (Interview 20).

Within the supply divisions examined the approach of officers was less concordant.[1] At the most general level—in rating the contribution of program evaluation to supply work—the responses from officers ranged widely both across *and* within divisions. At one extreme evaluation was assessed as constituting a "very small proportion of supply work" (Interview 10, 16). As will be discussed, judgments of this type are commonly linked to problems facing a number of supply areas in ensuring that the evaluation priorities, and hence information needs, of both portfolios and supply divisions are aligned. At the other extreme, evaluation in any form was always used in assessing programs (Interviews 06, 11, 14). At this end of the spectrum evaluation data was seen as making a "fundamental" contribution to decision making, even though it seemed to add appreciably to the workload of supply officers.

In those cases where evaluation was seen as "pretty much enmeshed in supply areas" (Interview 14), evaluation was gauged to have been increasingly relied upon in one particular area—the preparation and review of savings options (SOs). SOs are briefing papers which identify government programs that may be of either reduced priority within the government's stated policy objectives or, alternatively, of diminishing cost effectiveness (DoF 1992c, p8). In those circumstances where the government desires expenditure restraint, SOs are designed to nominate possible options for reduced spending. Individual supply divisions within the Department of Finance remain the principal source of SOs,

[1] Interviews were conducted with senior officers (for the most part, branch heads) in a cross-section of branches within three supply divisions—Employment, Defence and Industry, and Social Security. The majority, however, were sourced from within Social Security. See Chapter One for more details on the limitations of methodology and data.

although the introduction of portfolio budgeting in the late 1980s has induced portfolios to produce their own options for program savings.

Quite obviously the information produced by portfolio evaluations is likely to be of some import in the identification of savings options. In theory, evaluations flowing from portfolio evaluation plans are directed at both how well major programs are achieving their objectives and, more incisively, the extent to which they remain relevant to government priorities and community needs. For the moment we can note that there has been little convergence between theory and practice. Certainly, according to the most receptive supply areas, evaluation results are being consulted in the preparation of savings options, but often this is more an attempt to obviate deficiencies in either evaluation planning processes or inadequacies in more general information sources (for savings options evaluation, even if it is deemed poor, is “usually the only type of information available” (Interview 11)). One example, from the industry branch of the Defence and Industry Division, illustrates this situation (Interview 06). Here officers start from the premise that they have a much better chance of building a “solid policy case” if evaluation data is referred to. This is further buttressed where the evaluation indicates that a program might be under performing (although it was conceded that this situation was “rare”). In order to increase their chances of making a successful case for savings, the branch will specifically direct their SO inquiries to that small population of programs where evaluations are *known* to have been conducted. During the 1994-95 budget round two SOs were prepared using this approach, with one leading to abolition of a program and the other resulting in a smaller funding base. In this case the limited number of evaluation results available to supply officers determined where SOs would be targeted, the unsatisfactory implication being that some central budgetary priorities are driven not by the policy preferences of government (or Finance) but by the evaluation preferences of individual portfolios.

Within this context the work of supply areas and the subsequent advice flowing to Cabinet is, to a degree, driven by portfolio agendas. The task of supply areas is to work within the limitations of the evaluation planning process so as to avoid this. This is easier in some branches than others. Take for example the industry branch referred to above. Many industry assistance programs have sunset clauses that impose a specified time limit on funding. In looking at any new policy or extensions to existing policy the branch can, with relative ease, determine the budget subsets that the portfolio can pursue for renewal. In these cases the branch will track evaluations of particular programs and attain an adequate level of priorities coincidence (Interview 06). In other branches, this kind of evaluation tracking is more difficult because of the long lead times for programs to take effect. This is most pronounced in the health and social services sectors (Interview 14, 16).

In those cases where evaluation results are available, there is a recognition that program evaluation is something of a two-edged sword for supply officers. Outwardly they must endorse the official position taken by the evaluation strategy that “the most effective defence of a good program is a sound evaluation” (DoF 1992a, p7), but there is also a presumption in some supply areas that an evaluation is not a sufficient defence of a program since it would constitute irresponsible practice on the part of supply areas to guarantee the security of a program on the basis solely of an evaluation (Interview 05). At the same time, in other supply areas the position seems to be that, if anything, evaluation results that affirm a program are probably the best way to avoid arbitrary cuts—the problem in Finance’s management of evaluation is what, to line departments, appears as the “continuing capriciousness” of supply divisions in picking off programs one by one (Interview 06, 14).

There seems, therefore, to be some evidence of an evaluation disjuncture operating at the central budgetary level—a non-alignment of evaluation priorities between the central budget agency and line departments. The problem, however, appears to be a continuing one. As part of its efficiency audit of the evaluation strategy, the Audit Office examined the use of evaluation in the federal budget round of 1990-91, arriving at similar conclusions (ANAO 1991b). The audit, which was based on both internal Finance analysis and independent audit findings, noted that “in the majority of green briefs, savings proposals and new policy proposals, evaluation had not been mentioned” (ANAO 1991b, para 2.2.7, p14). This, of course, is likely to reflect the arrested flow of evaluations from guarded portfolios as much as resistance on the part of supply divisions. More tellingly, however, program evaluation was judged *not* to be a significant source of information contributing to budgetary analysis—evaluation priorities “were not reflecting government-wide information needs” (ANAO 1991b, para 2.2.21, p17). A crucial recommendation arising out of the audit was the “need for Finance, as part of the attempt to link evaluation to the budget process, to align more closely portfolio evaluation and its own savings agenda", links which "to date . . . have been neglected" (ANAO 1991b, paras 2.4.3-2.4.6, pp30-31).[2] The task for supply areas was to steer the primary instrument of portfolio evaluation—portfolio evaluation plans—towards the information needs of the central budget agency. It is to this process that we now turn.

Supply Influence Over Portfolio Evaluation Planning

There are three factors that combine to ensure that portfolios have the upper hand in setting the evaluation agenda, so that both the sequencing and substance of evaluations are not likely to coincide with either immediate or emerging budgetary priorities. First, supply division officers are not in a position to know most—let alone all—budget priorities in advance. The formulation of each year's budget is a protracted process in which changing political and economic priorities may quickly alter the information requirements of the Department of Finance's advisory role. Second, as we have seen, there is considerable variation in the use supply divisions make of evaluation in the preparation of advisory briefs for Cabinet. One important consequence of this is that the threat of having an evaluation used against a program is also variable. In the cases where they are aware of this situation, portfolios can use the evaluation planning instruments available to them to exploit this. This leads into the third factor. A central tenet of the government's evaluation strategy is that the conduct of evaluations and portfolio evaluation plans (PEPs) are the responsibility of individual portfolios (see Chapter Five). In this way PEPs, to the extent that participants take them seriously as a planning document, may be said to determine the evaluation agenda and hence the information flow to supply areas.

Certainly, a precursor to this problem had been recognised in successive efficiency audit reports where the Department of Finance had been chastised for its failure to enforce portfolio compliance with the planning requirements of the evaluation strategy (ANAO 1991a, para 5.1.13, p136; ANAO 1991b, para 2.4.31, p36). This lenient approach to evaluation planning represented what was perhaps the most conscientious application of devolutionary principles during

[2] It should be noted that, as a matter of course, the Department of Finance maintained that it has never had a savings agenda, "but rather produces savings papers as, increasingly, do all line portfolios, from time to time, to assist the government to meet its fiscal and program priorities" (ANAO 1991b, para 2.4.8, p32).

the long implementation period of the FMIP throughout the 1980s. One significant consequence of this perspective was that it seemed to distort the perceptions of those sections within Finance responsible for evaluation policy development as to the type of issues supply divisions were able to influence within the planning process. Indeed, according to the head of ESAB, supply areas "have a lot of influence on evaluation" and this could be exercised through a number of mechanisms (Interview 15; see also Mackay 1996, pp4-8). These included recommending which programs should be evaluated, determining the focus of particular evaluations—that is, the issues addressed and the questions asked—by influencing their terms of reference and participating on senior level evaluation steering committees. As a fail-safe, supply divisions were also in a position to brief the expenditure review committee (ERC) of Cabinet on the adequacy of evaluation in particular portfolios. The problem is that the recommended method bears little resemblance to execution.

Most branches of supply divisions concede having difficulty influencing PEPs. There are isolated examples of modest success, but this is most commonly contingency based. In the Employment branch, for example, supply officers set a target of sitting on approximately 30 to 40 per cent of steering committees established in the employment, education and training portfolio. The tight availability of time and staff resources, however, means that the branch must select what it regards as "the most important or high risk programs in terms of budget impact or new policy proposals" with considerable care (Interview 11). In a similar fashion, the industry branch also claimed to exercise some influence in determining priorities for evaluations, but also stressed that officers had to be alert in looking for "pressure points". These existed either where programs were approaching the end of their allotted funding, or where programs, although not high on the government's policy agenda, may be subject to trenchant criticism by the media or affected interest groups (and thus possessing potential to impact on policy agendas) (Interview 06). Still other supply division branches, such as health, tried to influence PEPs in more direct ways by either securing a Cabinet decision supporting the policy priorities as articulated by supply, or, in rare cases, using minister to minister liaison to overcome any recalcitrance in the PEP process (Interview 16, 18).

At the other end of the PEP spectrum were those supply division branches which claimed to exercise little if any influence over portfolio evaluation priorities. Apart from the fact that a minority denied that influencing evaluation was a "distinct component" of their work (Interview 06, 11), most of these accepted that "in theory" supply officers had the leverage to alter where evaluations were directed, but usually this was once the PEP draft was already in place. The health and general branch conceded that they had met with little success in this regard during the 1995 budget round. Further, some of the recommended mechanisms of influence were often turned against themselves—when, for instance, supply officers are invited onto steering committees by portfolios with the intention of "locking Finance into supporting or accepting the evaluation recommendations" (Interview 10). The capacity of supply areas to guide evaluation is also weakened markedly if there is no central evaluation committee or unit that is solely responsible for overseeing the conduct of evaluations within particular portfolios (a point confirmed by several audits of portfolio evaluation, see ANAO 1992a; ANAO 1993). In these cases, supply areas will often be dealing with individual program managers and hence open to being swamped by policy expertise and program detail (Interview 18).

The ambivalence displayed by Finance supply officers towards the PEP process does not come as a great surprise. It can, in part, be explained by the shape which the evaluation strategy finally assumed. There seems, at the highest level, to be some admission that for budget-related purposes the evaluation planning process has been a triumph of form over substance. In effect, PEPs (which are submitted by portfolios in February, although for the first few years of the evaluation strategy arrived in November) are difficult for supply areas to influence because they are not synchronised with the budget cycle (which culminates in the August Budget). This point was made by several supply branches (Interview 06, 10). As a consequence *systematic* evaluation planning will always be program management driven. This is compounded by the fact that Finance seems to have its supply hands tied—if Finance is to be seen to be faithful to devolved budgeting then, in a similar fashion, "supply authority to insist must be nil". In this respect the evaluation strategy was encumbered with conflicting if not unrealisable expectations, although wider, more oblique objectives clearly were well served:

> For evaluation to really feed into the budget then the intention for evaluation should be to coincide with two outyears. PEPs allow the government to have most of its cake and eat it too (Interview 25).

Portfolio Use Of Evaluation In The Budget

There is one last variable that influences the way in which supply divisions use evaluation in formulating their advice on budgetary policy—the extent to which portfolios are incorporating information from evaluations in their budget proposals. In this respect there is some validity in the argument that Finance budgetary advice will only be as good as the information provided by portfolios.

An instructive way of approaching this issue is to employ as a base for comparison the findings of independent reviews. The Auditor-General's series of efficiency audits provides an early assessment of portfolio evaluation use during the 1990-91 budget round (ANAO 1991b). After examining the information papers used in the preparation of the budget (access which no external researcher could hope to replicate), the audit reached some rather disconcerting conclusions. While guidelines for evaluation planning clearly required inclusion of "major evaluations" only, there was a general failure on the part of portfolios to restrict PEPs to this category. Instead, PEPs were collections of efficiency or process evaluations directed squarely at implementation issues rather than the results of program activity and their continuing relevance to government. Nor was there any indication that program effectiveness was being considered any more in 1990-91 than 1988, the first year of evaluation planning (ANAO 1991b, para 2.4.25, p35; para 3.2.5, p42; para 3.3.9, p47). Rather, PEPs were judged to have been consistently directed at lower level (or sub-program) evaluation planning rather than reflecting portfolio and government-wide priorities. The audit concluded that portfolios, concerned more with preventing program losses in the short-term, were giving insufficient attention to what evaluations might be relevant to future ERC (and hence Finance) consideration (ANAO 1991b, paras 1.2.16-17, p5; para 2.4.24, p35).

All of these findings found echoes in the views of supply division officers in 1995. There seemed to be three main complaints coming from supply branches. The first was that portfolios were simply not basing policy proposals on evaluation results, the efforts in this area often being described as "spotty" or "wretched" (Interview 10, 16). Instead, many policy proposals continue to be couched in terms of the outcomes to be expected rather than the outcomes that have already been achieved (and this despite pro-forma *Cabinet Handbook* requirements for all new policy proposals to indicate intended strategies for evaluation). In the experience of the health branch, "NPPs have been driven by new ideas, ministers and the community rather than evaluation" (Interview 16). In other words new policy was more likely to be driven by need—actual or ascribed—rather than mechanical, ritualised analysis.

The impression that portfolios were reluctant to use evaluation in budget submissions receives some support from an internal review of program evaluation within the (then) Department of Employment, Education and Training (DEET). DEET is one of the sprawling social service portfolios with a comparatively long history of *ad hoc* evaluation and review activity informing policy development. In terms of the central budget process, a sample survey of completed evaluations drawn from DEET evaluation plans between 1990 and 1992 found that only 10 per cent "were seen as having their main use in the Cabinet decision making context" and that "surprisingly little use" was made of evaluations to support Cabinet submissions (DEET 1994, p5). Although the survey asserted this was likely to be an underestimation because most respondents were in a better position to judge use in program management, it also conceded that utilisation in the budget process was judged to be lower than in other portfolios. One reason given for this low usage was that the

> frequent need to develop new proposals at short notice has complicated the task of having evaluation findings available to inform design of budget proposals (DEET 1994, p5).

This is not at all persuasive. A crucial objective of evaluation is to inform the policy development process, and the budget process is certainly part of this. The above reasoning definitely counters assertions by the Department of Finance and the Management Advisory Board that program evaluation should be seen as part of every day management responsibilities. It also neglects the fact that the pressures of short notice policy development are common to all departments, particularly central agencies. It is likely, however, that the situation in DEET could be attributed in part to the "somewhat threadbare" coverage of programs that the portfolio's PEP provided. Not only does this corroborate audit findings and interview impressions, it also suggests the possibility that the evaluation of major programs of a "sensitive" nature had been studiously avoided (DEET 1994, p9).

The second issue facing supply divisions related to those instances where evaluation *was* referred to in a portfolio budget submission. In a number of cases, supply division branches claimed that evaluation was being used, but used selectively—to *support* policy only (Interview 10, 11). This area is quite obviously the point of essential contestability between a budget agency and line departments; according to one supply officer, whose responsibility included the employment portfolio, that portfolio's interest in evaluation extended only as far as how to maintain programs rather than any other "higher" rationale, such as advising government on how to re-order its priorities (Interview 11). This view is also subscribed to in some line portfolios which claim that the evaluation planning process encourages weak policy analysis that reaffirms for departments the general effectiveness of programs rather than pinpointing problem areas that could be addressed:

> The evaluation process tends to produce outcomes supportive of programs, often indicating that more output could be achieved if more resources were made available. The process does not seem to deliver the best input to the budget priority process, in an environment of declining resources, particularly when something(s) need to be cut if high priority new initiatives are to be accommodated (DPIE cited in MAB-MIAC 1993b, p468; similar reservations can be found at DEET 1994, p12)

As this account suggests, program evaluation was, for some portfolios at least, being used (or misused, from the perspective of Finance) solely as a defensive mechanism—the planning of evaluation and its results either reinforced a predilection towards aiding programs or were manipulated so as to confound rather than inform the budgetary process, for example through the placement of "minor" evaluations on PEPs.

This brings us to the third issue identified by supply officers concerning the alignment of information priorities—portfolios, through both portfolio planning and budget submissions, simply were not delivering the kind of performance information that Finance *could* use for its review obligations (Interview 10, 11, 14, 16, 18). 'Could' is the operative word here, since it should be recalled that while the twin dilemma of non-use or selected use of evaluation was the most visible manifestation of the dissymmetry of evaluation planning, the demand for information relating to the effectiveness and relevance of government programs demonstrated high variability across different supply divisions. The problem was widespread and persistent, although there were late signs of change as the case of the Department of Human Services and Health (DHSH) illustrates.

The HSH portfolio, a product of the machinery of government changes in July 1987, federated a number of agencies with disparate policy concerns as well as embedded and competing professional perspectives. The portfolio's PEP reflected this, surveying a great deal of sub-program level review activity that was inadequately related to questions about the effectiveness and relative importance of different programs. Indeed, the PEP was described by the portfolio as not a "true plan" but rather a "list" (DHSH 1994, piii). These limitations had long been suspected within Finance, but, given the central agency's posture on devolved evaluation responsibility, it was not until an internal review in 1995 that the full extent of the health portfolio's oscitation on program evaluation was revealed.

The review, which was seen by supply officers responsible for health to be something of a "rescue attempt", made some remarkable findings. The term 'program' had been interpreted by the portfolio "at a much lower level than intended"—embracing a hotch-potch of reviews that were effectively a "substitute for ongoing management activities"—with the consequence that of 181 evaluations listed in health PEPs *none* were at the program level. This situation was exacerbated by two factors. First, there occurred a further "misinterpretation" of Cabinet Handbook requirements for each NPP to be accompanied by a tailored strategy for evaluation that encouraged a "cover-the-field" approach to planning evaluations. Second, after some seven years of the evaluation strategy, the DHSH had yet to establish a central oversight mechanism for selecting priority programs for evaluation and monitoring their progress (Amies 1995, pp1-2). It was not until mid 1995 that a Portfolio Strategies Group (PSG) was, with the blessing of Finance, finally established in the health portfolio and assigned responsibility for identifying "fewer, key strategic evaluations" designed specifically to assess program and cross-program results and their appropriateness to government policy objectives (Amies 1995, p4; DHSH 1995a, pp1-2; DHSH 1995b, pp5-6).[3]

[3] This about-face within the health portfolio can be attributed to changes in key personnel. After a procession of secretaries in the late 1980s and early 1990s, each with widely varying commitment to the government's evaluation strategy, it was at this time that a semblance of stability emerged with the appointment of Dr Stephen

In summary one might argue that, in terms of the Department of Finance's long and much publicised aspiration for the integration of program evaluation into budgetary decision making, the evaluation strategy has met with some insurmountable obstacles. These hurdles are to be found on both sides of the budgetary divide. The most crucial observation to make at this stage is that the attempt to firmly link evaluation to the budget has collapsed into an almost self-defeating posture. Clearly, in institutionalising evaluation Finance sought to increase the receptivity of line departments to evaluation by establishing it as their responsibility. The danger is that the process is heavily weighted *against* the information needs of the central budget agency, since it permits line departments to both determine the evaluation agenda and unduly influence budgetary and policy priorities via the PEP process.

This initial assessment does, of course, presume that both Finance and line departments are sincere in their approach to the value of the PEP process. There was some evidence that for practically the entire period that the PEP process was in operation both Finance and line departments treated it as predominantly a procedural obligation that, at best, *might* supplement either Finance's capacity to direct evaluation activity or strategic development by portfolios. (In this respect the secretary to the DHSH was particularly scathing, claiming that "evaluation is done only to meet Finance requirements" with the knowledge that "PEPs are irrelevant" (Interview 07)). Hence, the dangers of the process were akin to a candle burning at both ends. At one end a number of portfolios, safe in the knowledge that Finance could not or would not impose evaluation priorities to align with projected budget imperatives, would engage in budgetary machinations (by way of blatant avoidance or bureaucratic ritual). At the other, formal evaluation obligations have had an equivocal impact on Finance supply division work practices. Despite the fact that central agency control is strongest in informal power networks, if supply divisions are reluctant to adopt, or resist integrating, evaluation into their review procedures then the thrust of formal evaluation planning is severely blunted.

Duckett as secretary to the portfolio. Duckett, an academic health economist with substantial experience of program evaluation, initiated major divisional restructuring and forced the evaluation issue with senior management in the department (*The Canberra Times*, 18 June 1995, p10). A further impetus came with the transfer of Dr Marion Amies, an evaluation specialist from ESAB in the Department of Finance, to the health portfolio in early 1995.

Of course, there is no doubt that in some portfolios, and in specific instances, evaluation *is* informing aspects of the budgetary process (most notably mechanisms for identifying savings). From earlier discussion the Industry and Defence Division, which already had a predisposition to evaluation and review activity, is a good example. But the use of evaluation is, on a portfolio basis, often dependent on prevailing circumstances and perilously close to occurring on an *ad hoc* basis. It almost goes without saying that an outcome of this nature counters one of the primary intentions of the evaluation strategy, which was to ensure systematic planning, conduct and use of evaluation to inform central budgetary decision making. The Audit Office cautioned in its 1991 efficiency audit that:

> the program evaluation initiatives had not yet produced the significant leap forward sought by the Government in improving the level of program performance information available for use in the central Budget context (ANAO 1991b, para 1.3.1, p6).

While conceding that there would always be limits on both evaluation's use as a "definitive solution" (ANAO 1991b, para 2.1.14, p6) and external assessor's capacities to measure its utilisation, in essence the audit report concluded that the central budgetary process was not being appropriately informed by evaluation. Based on interviews with senior supply officers, this present examination is approaching a similar conclusion. But first we should compare this admittedly impressionistic interpretation with some quantitative research undertaken by the Department of Finance that sought to demonstrate linkages between evaluation and budget processes.

PROGRAM EVALUATION AND THE BUDGETARY PROCESS: SIFTING THE EVIDENCE

To this point an argument has been framed around reservations relating to two developments in evaluation and resource coordination. First, that policy evaluation has long been a part of the armoury of central budget agencies in Australia and that the evaluation strategy does not necessarily represent a radical departure from the much maligned conventional practices. Second, that the mechanism for tying program evaluation to the central budgetary process—portfolio evaluation planning augmented by supply division intervention—was not living up to the government's expectations. In order to sustain these contentions two types of additional evidence will be presented: an analysis of data collated by the Department of Finance relating to the use of evaluation in the budget, and the move towards Cabinet commissioned program reviews as a substitute for program evaluation in advising on budgetary priorities. Both sets of evidence indicate that in terms of assisting policy advice on budgetary matters the evaluation strategy was having only a marginal impact.

Mixed Signals: Finance Monitoring Of Budgetary Use

The Department of Finance places great store on its almost unique position to assess the contribution of program evaluation to decision making by Cabinet, and, in particular, its expenditure review committee. One officer has commended the process as follows:

> Our people sit in on Cabinet deliberations. We hear the types of comments made by Cabinet ministers and we hear the level of emphasis put on evaluation information being available to help them in their decision making. We hear their frustration and dissatisfaction when adequate evaluation information is not available—for example, when it has not been completed in a timely manner (Mackay cited in SSCFPA 1994, p189).

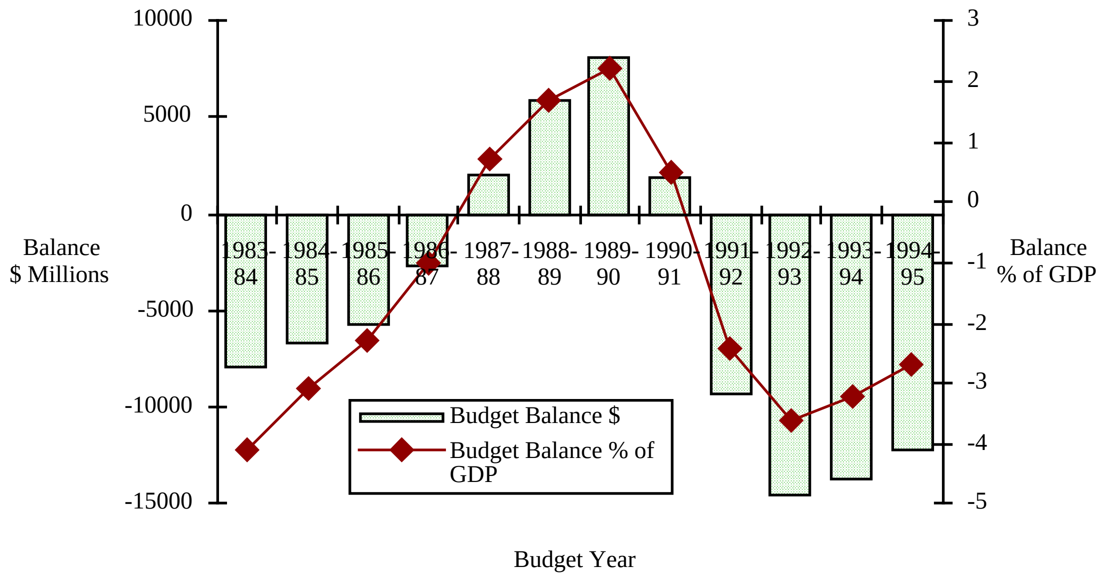
This description makes reference to two significant points. First, when the review process relates to *cabinet* decision making, it will rely on the judgment of Finance officers who attend budget deliberations. It is, in the end, a highly informal assessment method. Second, it concedes that the evaluation strategy may not be satisfying the advice requirements at the peak decision making level of cabinet. This provokes a number of questions. Do ministers, especially non-finance or spending ministers, give much credence to evaluation when deciding program priorities? To what extent *is* evaluation incorporated into ERC advice sourced from Finance? Given the confidential nature of cabinet decision making, can these issues ever be resolved adequately? The Department of Finance has certainly made a start in its *Evaluation in the Budget* series which presents information on the utilisation of evaluation in budget submissions by Finance supply divisions and portfolios, as well as an indication of subsequent use by the ERC of Cabinet (DoF 1992c, 1993b, 1994b, 1994c).

In analysing this data the intention here is not to enter the domain of ministerial decision making, although this runs the obvious risk of omitting the key test of verification. Certainly, the Finance papers do try to gauge this and what little is available will be considered. But there are a number of factors that constrain a focus at this level. One is the difficulty of attaining the necessary data from former cabinet members. Another relates to the problems of explaining movements in budgetary policy at the aggregate level. As discussed in earlier chapters, the Labor Government operated within a period of acute fiscal stress. The key issue here is separating the effects of the political will displayed by cabinet ministers in restraining expenditure and the contributory impacts of new resource coordination systems (such as the changes to forward estimates and the evaluation strategy) (see MAB-MIAC 1993a, pp226-230). There is some suggestion that the introduction of a more tightly controlled forward estimates system, in conjunction with deteriorating economic conditions, permitted more strategic ministerial consideration of policy proposals:

> Previously, Cabinet devoted considerable effort to establishing the baseline for each budget. This was because agencies were required to bid for resources to carry on existing policy each year. The introduction of the forward estimates system has freed Cabinet of that task, allowing it to concentrate more precisely on the strategic focus of the budget. Accordingly the content of policy advice is now more oriented to broader outcomes issues and to overall budgetary processes and objectives to achieve them (MAB-MIAC 1993a, p226).

The contention is that by forcing ministers to consider the future consequences of spending decisions and curtailing the discretion that departments had over bids for future years, the forward estimates process impacted on budgetary decision making. Certainly, the 1993 MAB-MIAC evaluation of public service reform was quick to point to “a progressive reduction in the rate of growth of forward estimates since 1981-82” (MAB-MIAC 1993a, p228). This fiscal discipline is illustrated by the budget outcomes of the Labor Government between 1985 and 1990, although obviously less so between 1991 and 1995 as policy priorities shifted under the leadership of Prime Minister Keating (see **Figure 6.1** below).

Figure 6.1: Commonwealth Budget Balances, 1983-84 to 1994-95



Note: Estimate for 1994-95

Source: Commonwealth of Australia 1995, p1.11.

For the MAB-MIAC evaluation of public service reform the impact of the forward estimates on cabinet budget deliberations was clear-cut—it had an unequivocally positive impact (MAB-MIAC 1993a, p230). It is doubtful, however, whether this same conclusion could be applied to the influence of program evaluation on cabinet decision making.

Nonetheless, if program evaluation was to have a decisive impact on the budget process it was in the period *after* 1990. There are two reasons for this. First, by this time the evaluation strategy planning cycle had been operating for some four years. A reasonable expectation might be that most implementation problems would have been addressed, although, as the previous section described, anecdotal evidence suggests that Finance use of evaluation was irregular as late as 1995. Second, the government's fiscal policy had shifted towards an expansionary posture. As a percentage of Gross Domestic Product (GDP) the status of the federal budget was to move from a 2.2 per cent surplus in 1989-90 to a deficit high of almost 3.6 per cent in 1992-93 (see **Figure 6.1** above). The aim of this policy shift was to direct increased expenditure at several priority areas in order to alleviate the worst effects of the recession of the early 1990s; for example, the *One Nation* statement of February 1992 increased funding significantly for economic infrastructure consolidation and family allowance payments (see Keating 1992, pp4-9). This spending was financed from two sources—revenue measures (principally the sale of public assets) and trading off savings in existing programs for new policy. However, it must be noted that in the three budget rounds after 1992-93 the government also set itself the objective of reducing the budget deficit to 1 per cent of GDP by 1996-97 (Commonwealth of Australia 1993a, 1993b). In both periods the logic of the evaluation strategy dictates that program evaluation *should* play a role in providing effectiveness information for these program trade-offs.

The internal survey series conducted by ESAB represents the only publicly available 'hard' or quantitative evidence relating to evaluation usage in the federal budget. We know that the surveys have standing within the Department of Finance, since a number of senior officers have employed the findings in public statements on the evaluation strategy (eg Sedgwick 1993; Mackay 1993, 1994). The surveys are based on questionnaires directed to the various supply divisions in Finance and, in the case of Cabinet decision making, those officers who attended ERC deliberations (DoF 1993b, pp1-2, 15-16; questionnaire appears at Appendix III). The information they contain is racked by ambiguity, but is here interpreted as supporting the argument that evaluation has failed to either fundamentally recast the work of supply areas or increase the impact of analysis in ministerial decision making. This conclusion rests as much on methodological inadequacies as the substance of findings, although as we will see what is often of most interest is what the surveys fail to discuss.[4]

---

[4] Over the course of the survey the methodology employed and, perhaps more significantly the variables measured, changed markedly. This makes comparison between years and variables extremely hazardous (and serves also to undermine the general trends identified by the series). The methodological problems include, but are not confined to, the following:

(a) The survey gradually confined its analysis to a sample of policy proposals, moving from an examination of all proposals (1990-91, 1991-92) to a large sample (1992-93) and finally a small sample weighted to large proposals (1993-94, 1994-95). On this basis the reliability of the survey *from year to year* could be questioned;

(b) There is an unsatisfactory discontinuity of measurement for a number of key variables: (i) information on green briefs is included in the first survey but abandoned for the remainder of the series; (ii) the classification of savings options according to either portfolio or Finance preparation is included for only one year (1992-93); (iii) unsatisfactory discontinuities on information on Cabinet decisions relating to savings options influenced by evaluation, and (iv) incompleteness on Cabinet decisions on new policy proposals which were jettisoned after the 1993-94 survey;

At the most general level the surveys concluded that “the extensive use of evaluation findings has become an established feature of the budget process in the Commonwealth Government” (DoF 1994b, p1). Use is assessed by supply officers according to the type of influence that evaluation—as defined within the scope of the evaluation strategy—has had on a particular proposal. This could take one of two forms. The influence of evaluation was judged to be *direct* when “a proposal resulted directly from the findings of an evaluation”; alternatively influence was *indirect* where a “proposal uses evaluation results either as a source of background material or the evaluation helped to create a persuasive climate of opinion” (DoF 1992c, pp5-7; DoF 1994c, p2). The indirect influence classification was intended to encompass the cumulative impact of evaluation on policy development.

As described earlier, new policy proposals (NPPs), which are sourced almost solely from portfolios, and savings options (SOs), which are prepared by both Finance and portfolios, are the object of most bilateral budgetary negotiations between supply divisions and portfolios within the portfolio budgeting framework. **Figures 6.2** and **6.3** show the use of program evaluation to support NPPs and SOs considered by the ERC of cabinet for the five budget rounds covered by the survey. The charts indicate that both NPPs and SOs have been significantly “influenced” by program evaluation. For NPPs the influence has most consistently been direct, peaking in the 1994-95 budget round at 77 per cent of the total value of proposals. For SOs the type of influence has varied appreciably between direct and indirect, peaking during the 1991-92 budget round at around 68 per cent. Although the surveys have claimed that the use of evaluation in budgetary proposals has “trended upwards”, it is perhaps more accurate to conclude, as an internal survey did, that “the [survey] trend has been erratic and the credibility of the information dubious” (DoF 1995, p7). We can note a number of points that frustrate the kinds of results reported in the charts. In line with the analysis so far, these occur at two levels—the impact on

---

(c) The analysis of evaluation use by cabinet is problematic since it is based on judgment by a small number of Finance officers who attended ERC deliberations. There is considerable difficulty in assessing “decisions” in the cabinet context and what the basis was for any particular decision—there is likely to be a good deal of “second-guessing” of cabinet’s decisions. This is exacerbated by the subsequent disclosure that a “majority” of senior supply officers considered the survey questions unclear and inappropriate (DoF 1995, p6).

the preparation of budget submissions from both Finance and portfolios, and the effect on cabinet decision making.

Budgetary Processes

In so far as the data is available (see footnote 4) the surveys corroborate the findings of the 1991 efficiency audit of evaluation in the budget—supply divisions in Finance have no greater propensity to utilise the results of program evaluations, particularly in the preparation of SOs. One stark indication of this is that in the *only* survey to divide the preparation of SOs between portfolios, it was found that less than 2 per cent of Finance prepared savings options were in fact influenced by evaluations (DoF 1993b, p6). Since "in dollar terms most major SOs are initiated by Finance" (DoF 1994b, p4) one could surmise that a very large proportion of SOs are prepared by Finance without the aid of information derived from evaluation, let alone other performance information relating to effectiveness. Similarly, in the sole survey to examine the use of evaluation in Finance-prepared "green briefs" (which are intended admittedly to complement rather than compete with NPPs and SOs) only 33 per cent referred to an evaluation of any type (DoF 1992c, p15). In both instances the reasons as to why both sets of information were discontinued is not provided, and although our interviews revealed that supply areas were increasingly relying on evaluation in the preparation of SOs, the conspicuous failure to report sends signals of doubt.

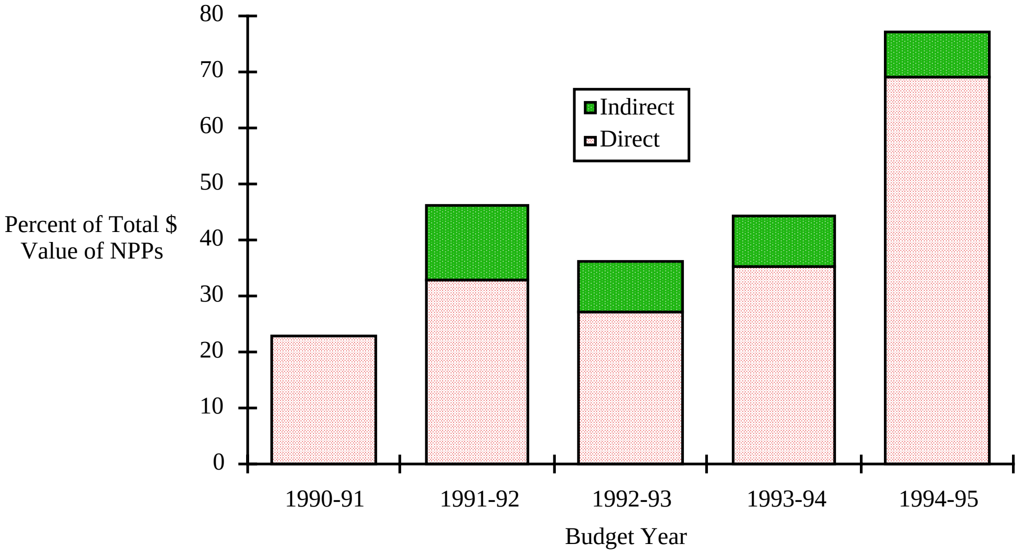
Cabinet Decision Making

It is not clear from the survey results whether evaluation influences cabinet decision making. On the one hand the data suggests, at a broad level, that Finance officers judged evaluation to have some impact on cabinet decision making. **Figure 6.4** below shows, where the data is available, the proportions of cabinet decisions taken on NPPs and SOs assessed by Finance officers who attended ERC meetings to have been influenced by evaluation. This influence was extremely variable, most noticeably for SOs. Importantly, the survey fails to consistently compare the impact of those proposals supported by evaluation with those that were not. The most revealing results come from what the survey excludes.

First, the use of evaluation in the preparation of SOs is wildly variable, not only in aggregate terms but also in the split between direct and indirect influence. Of most significance is that of the SOs influenced by evaluation only a small proportion were judged to have influenced Cabinet decisions in 1992-93 (26 percent) and 1993-94 (16 percent), although this increased to an astonishing 94 per cent in 1994-95. Similarly, in the two rounds in which information on the impact on Cabinet decision making of those NPPs supported by evaluation and those that were not were compared, the survey found that NPPs supported by evaluation had practically the same chance of being accepted by Cabinet as proposals that were not supported by evaluation—74 as opposed to 76 percent in 1992-93, and 71 as opposed to 72 percent in 1993-94 (DoF 1993b, p9; DoF 1994b, p5).

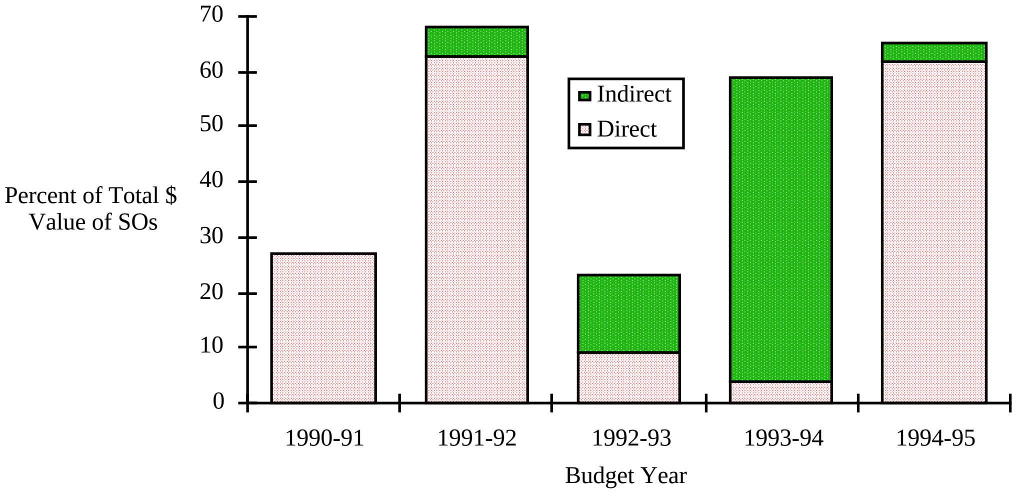
Despite the variability of these figures, one interpretation might be that the effectiveness of the policy advice to Cabinet produced by both Finance and portfolios is not significantly increased by evaluation when it is available. This is borne out by an internal Finance review of the *Evaluation in the Budget* series which cited a number of Finance officers with experience in Cabinet to the effect that “evaluation was not often discussed in the Cabinet room during consideration of NPPs and SOs, unless it was interdepartmental reviews instigated by Cabinet themselves” (DoF 1995, p5). However, in other quarters the one benefit that evaluation was seen to bring to the cabinet room was that it obliged ministers to “look inside a proposal” rather than simply concentrating on the “proposal’s label and dollars” (Interview 06).

Figure 6.2: New Policy Proposals (NPPs) Influenced by Evaluation in Budgets, 1990-91 to 1994-95



Source: DoF 1994b, p3.

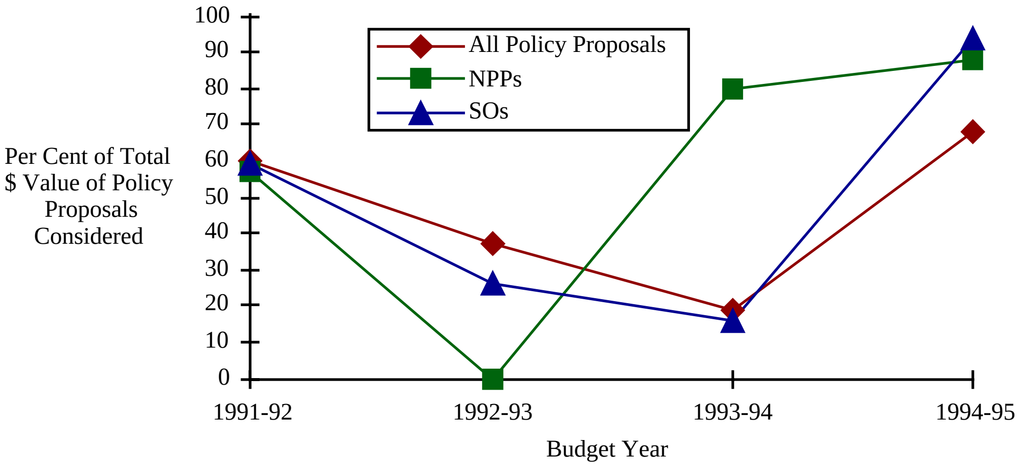
Figure 6.3: Savings Options (SOs) Influenced by Evaluation in Budgets, 1990-91 to 1994-95



Note: Direct and indirect influences not distinguished for 1990-91. Included as 'direct'.

Source: DoF 1994b, pp4-5.

Figure 6.4: Influence of Evaluation on Cabinet Decisions in Budgets, 1991-92 to 1994-95



Note: Data on NPPs for 1992-93 not available.

Source: DoF 1992c, 1993b, 1994b, 1994c

Of course, Finance rationalises the survey results by arguing that the inconsistent use of evaluation "depends crucially on the nature of the particular budget environment" and that "unexpected policy changes meant there was less scope for formal evaluation, with its lead times, to be influential" (DoF 1992c, ppii, 1; 1994b, p1). Further, it contends that "each budget is strongly influenced by the Government's policy agenda and urgent economic and other imperatives, the relative importance of which vary considerably from year to year" (DoF 1994c, p1). This position is untenable on two grounds. First, the budget *is* the government's policy agenda. The envisaged role of systematic evaluation is to improve the advice which assists the government in determining its policy agenda. In effect, the assertions above concede that program evaluation is unfit for the task of informing budgetary decision making because of timeliness and relevance issues when policy priorities change. Second, to assert that both policy conditions and the priorities of the budget can and do change is to state the obvious. This reason is not persuasive precisely because the rationale driving the evaluation strategy was to aid decision makers—both political and bureaucratic—when altered policy conditions called for a reordering of objectives.

However, the defence outlined by Finance does serve one purpose—it suggests that the singularity of some policy decisions in certain budgets does explain a lot about the level of evaluation utilisation reported in the surveys. The 1992-93 budget round, for instance, was framed by worsening economic conditions that provoked the hasty development of a comprehensive range of employment programs, including the Labor Government's "Employment Package" of labour market programs and the *One Nation* statement that committed new funding to transport and communications infrastructure measures. These programs were principally the product of extra-budgetary policy development. Similarly, the 1994-95 budget round was dominated by the government's *Working Nation* initiatives which were heavily dependent on evaluations of employment and industry policies. In that year *Working Nation* accounted for some 60 percent of the total value of NPPs. From **Figure 6.2** above, for NPPs that did form part of *Working Nation*, the proportion of NPPs judged to have been influenced by evaluation falls from some 77 to 58 percent (DoF 1994c, p3). These figures still indicate a significant level of evaluation utilisation, but in both cases the figures were inflated because of a single major policy development which consumed resources. If the impact of the evaluation strategy is measurable only against specific budgetary environments then the systematic application of evaluation is undercut.

On any reasonable interpretation, this analysis of the survey's evidence has not been able to demonstrate *conclusively* the ineffectiveness of the evaluation strategy for central budgetary decision making—methodological inadequacies and highly selective reporting of data will preclude this. Nor was establishing such failure the desired intention. Quite obviously, and allowing for the limitations of the information, evaluation is being used to *some* degree at both bureaucratic and political levels.[5] The analysis has, however, raised a sufficient number of reservations about the survey analysis to indicate that the strategy is not realising its ambitions for the systematic flow of evaluation into budgetary decision making or the advice that informs it.

---

[5] Although the survey is silent on the type of evaluation used. That is, were the evaluations used dealing with questions of program efficiency or the more crucial questions of effectiveness and appropriateness? The survey itself was a response to criticism contained in the first efficiency audit of the evaluation strategy that Finance had not developed performance indicators for the strategy (ANAO 1991a, para 4.7.20, p120). At its base, however, the survey was a "crude" process or quantity measure rather than a gauge of the quality of evaluation in the budget (see DoF 1995, pp4, 10). The survey was discontinued after the 1994-95 budget round.

Topping Up? Budget Related Reviews And FESPs

Earlier discussion referred to the return on the part of the Labor Government to an economic policy stance of fiscal contraction after the 1993 federal election. Policy developments such as those contained in *Working Nation* signalled that expenditure restraint was to be pursued within the framework provided by the government's explicit social justice objectives. The evaluation strategy, however, was not the sole preferred mechanism. New options were required for containing expenditure growth in order to meet the government's medium term economic objective of reducing the budget deficit to 1 percent of GDP by 1996-97 (Commonwealth of Australia 1993a, p24). The government's response was to look at two alternative advisory mechanisms—the commissioning by cabinet's ERC of a wide-ranging schedule of "budget related" program reviews and the preparation by the Department of Finance of forward estimates strategy papers (FESPs) that where intended to advise cabinet on the sustainability of specified program expenditures. It is significant that the FESPs process *as well* as the program review schedule were *Finance initiated.* The contention here is that these developments are further indication of the failure of the evaluation strategy adequately to inform budgetary policy advising and decision making.

Although the ERC first requested "program reviews" be prepared for the 1990-91 budget round (ANAO 1991b, para 2.4.2, pp30-31), they were subsequently commissioned on a more expansive scale as part of the 1993-94 budget round in order to meet the government's announced deficit reduction commitments. The Government outlined the rationale for the reviews as follows:

> taking account of an analysis of outlays trends across all portfolios together with Minister's own reviews of their portfolio priorities, the Government is putting in train a program of reviews, extending the Government's normal program evaluation activity. The reviews will focus primarily on areas which have been subject to rapid outlays growth in recent years or which for other reasons are seen as offering scope for program improvements (Commonwealth of Australia 1993b, p3.13).

As this states the reviews were seen by the government to be an extension of "normal program evaluation activity". The schedule of program reviews was extensive, having a three year horizon framed around the budget cycle, and covering programs as diverse as significant outlays expansion in the health system, funding for the Australian Broadcasting Corporation, natural resources management programs and the appropriateness of links between the age pension and superannuation entitlements (see Commonwealth of Australia 1993b, pp3.13-14; Commonwealth of Australia 1994, pp3.11-13).

However, there seemed to be competing purposes for the reviews. At one level the government sought very specific analysis designed “to assist in formulating strategies for containing outlays growth” (Commonwealth of Australia 1994, p3.11). At another, senior Finance officers were extolling the reviews as much more than an exercise in achieving aggregate targets for expenditure reductions, as the following two comments convey:

> . . . the reviews were not designed solely to look for savings. We are primarily interested in the effectiveness of programs. In some cases where you can see a way of doing something better there might be savings. Alternatively, there might be a way within the existing allocation of funds of just getting a better result (CPD, Senate, Estimates Committee D, 2 September 1993, pD249).

> Perhaps the best way to regard reviews is as being complementary to the normal cycle of PEP evaluations, rather than being a substitute for them or reflecting any deficiency in the system of PEP evaluations. The former allow an urgent response to emerging budget pressures, while the latter allow a more measured and progressive coverage of programs and of important program issues (Mackay 1994b, p237).

These two objectives do not sit together comfortably. If the reviews were merely an augmentation of program evaluation, why would they be examining the kinds of effectiveness and appropriateness issues that, presumably, had been covered by PEP program evaluations? Furthermore, as the second extract suggests, it is stretching credulity to argue that program evaluation provides a “more measured and progressive coverage of programs” if, as the previous two sections have argued, that analysis is either not forthcoming or not referred to by supply divisions when advising on budget submissions. More to the point, as already discussed, the entire PEP process was predicated on the capacity of supply divisions to tailor evaluation planning (and hence the questions asked by program evaluation) to emerging budgetary priorities. To the extent that the evidence presented earlier suggests that this is not occurring, then the shift towards program reviews does represent a "deficiency" in aspects of the evaluation strategy.

This interpretation is strongly supported by supply division officers who were assigned primary responsibility for undertaking program reviews. In the main, reviews were seen as a response by the government to the constraints of a budget process that squeezes out detailed analysis—this could be considered one of the defects of the budgetary end of the evaluation strategy. As a consequence, a number of supply divisions claimed that ERC commissioned reviews had, since 1993-94, emerged as the staple of evaluation activity upon which budgetary advice was based; they "had more impact" and their urgency took precedence over evaluation planning processes (Interview 10). In certain branches, especially health and social security, program reviews were judged to have consumed vast amounts of time and staff resources, estimated at something like 60-70 per cent in some cases (Interview 14, 16, 18).

Program reviews generally adhered to a common methodology. Reviews were commissioned by ERC based on Department of Finance advice on their need. Three principal criteria were considered—perceptions of escalating cost, evidence of under performance, and (or alternatively) a lack of sound information on which to make judgments about performance (Interview 14). Portfolios were given an opportunity to make separate submissions on reviews to ERC, but it is more than likely that Finance proposals generally prevailed. One of the instruments introduced contiguously to support the reviews were forward estimates strategy papers (FESPs). As we have seen, program reviews were directed at a specific target—what was perceived to be unsustainable growth in some program outlays which narrowed the scope for policy trade-offs. The role of FESPs was to "better structure" outlays growth analysis by tracing expenditure trends, isolating major underlying causes and advising on policy issues that could impact on the budget in future years (Commonwealth of Australia 1993a, p24; Glenn 1994, p7). FESPs and hence program reviews were "not evaluations, but evaluative activity". Their task was to save money by restricting potential outlays growth without reducing program effectiveness, and in this they placed the government's economic (as opposed to social) objectives in relief (Interview 16; Glenn 1994, pp10, 16). Hence, reviews were not directed at fundamental questioning of programs.

An illustrative example comes from the health branch. The Medical Benefits Scheme (MBS), which provides for a schedule fee rebate for medical services, is the central plank of the public health system in Australia known as Medicare. Costing approximately $6 billion annually but increasing by around 5 per cent each year, it was identified by Finance as one of the fastest growing components of outlays. A close analysis of contributory factors was deemed necessary, but there was found to be neither sufficient time in the budget process nor adequate portfolio program evaluation addressing the issue. In the end, Finance drew on existing information on Medicare held by the Health Insurance Commission.

Commissioned during the 1993-94 budget round it was intended that the "pay off" from the review would come in the 1994-95 round. However, the MBS as a whole was not subject to review, only components. The health branch proceeded from the assumption that Medicare "worked":

> the appropriateness of programs such as Medicare and Child Care which were seen as part of the landscape, were therefore not looked at as fundamental policy. The central parameters were not questioned; we were working at the margins (Interview 16).

This was reflected in the reaction of Cabinet during the 1994-95 Budget to the branch's co-payment proposal for general practitioner visits under Medicare. This was rejected due to the philosophical commitment of the government to universal free health care under Medicare. In the end two, more palatable proposals were adopted—reducing payments for pathology services and curtailing supplier induced demand for health services by reducing the number of medical graduates.

Program reviews signalled a return by the Department of Finance to its traditional concerns with expenditure restraint since, in terms of evaluation priorities, they wrested control of the budgetary agenda away from portfolios. To some extent they were also a sharp reflection of the altered policy objectives of the Labor Government towards the end of its final term in office—after considerably freer spending in the early 1990s, fiscal responsibility once again entered the government's policy agenda as a political imperative. One perspective of the program review process was that the government took a risk in commissioning the reviews since by lifting some programs from its normal evaluation processes it may have caused others to be taken off the policy agenda (Interview 25). Another, just as plausible view might be that the government was transferring programs from the evaluation planning process *on to* its budgetary agenda.

To conclude, the fact that program reviews were commissioned on a regular basis after 1993-94 indicates either a growing dissatisfaction on the part of Cabinet's ERC and their (ministerial office) advisers for the products of the evaluation strategy, or a reflection of advice from senior Finance officers for a review mechanism more fully integrated with the budgetary process and, importantly, more amenable to *direct Finance control.* Reviews indicated a return to the more basal concerns of traditional expenditure restraint. We would do well to remember that for all the adherence to the form of Cabinet endorsement, in substance the evaluation strategy was essentially a bureaucratic—and in particular, Finance—led initiative. Program reviews appear to be an admission by the government that the evaluation strategy was not producing the kind of information required by the ERC at the time it was required. A secretary to the Department of Finance conceded as much, claiming that ministers simply "weren't getting the right information at the right time" and, given the propulsion of the PEP process by portfolios, "we [the Department of Finance] wanted the government to be in a position within the budget cycle to look at what was driving budget outlays" (Interview 25).

CONCLUSION: RUNNING TO STAND STILL IN RESOURCE COORDINATION?

It is fair to say that both the Hawke and Keating Labor Governments and the Department of Finance asked a great deal of the evaluation strategy. Program evaluation was introduced as a key component of financial management reforms designed to emphasise the results of government activities. It was also grasped as one way to modulate the government's somewhat sporadic commitment to what has been termed "decremental budgeting" (Schick 1986b, pp290-297). In this situation, as we have seen, fiscal stress compels a government to make explicit decisions about resource allocation where in order to redistribute benefits to one group they must necessarily be taken from another. When this redistribution bites into the base budget of specific programs, this type of budgeting can heighten dangers of increased interest group conflict and decreased budgetary stability. The evaluation strategy was directed at recasting the budgetary norms so that relative priorities based on systematic analysis of program effectiveness became a prime consideration in this type of decision making. A common information base for the resource coordination process would palliate the conflict of (departmental) interests and ultimately aid the ministerial task of making the best use of limited budgetary resources.

As a consequence, this chapter set out to assess the impact of the government's evaluation strategy on the resource coordination dimension of central agency policy management capacities. It was suggested in an earlier chapter that this impact could be gauged by focusing on how aspects of the budgetary process are now conducted. Two sets of evidence were presented. The first, based on interviews with budget agency participants, suggested the impact of program evaluation on the work of supply divisions has been marginal. As an instrument for advising both on the merits of policy submissions and the location of possible savings measures, the evaluation strategy has proved difficult to manage—the evaluation planning process was, by design, heavily weighted against the central budget agency; the available performance information was either not relevant to cabinet level priority setting or was deployed by portfolios to act as a shield for programs; and the evaluation planning cycle seems not to fit into the overriding budget cycle. Just as significantly, at the time of interview, there remain significant pockets of scepticism and resistance on the part of supply areas towards evaluation. To these officers *all* program information is vital in their task of reviewing budget proposals, regardless of its source.

The second body of evidence, based on utilisation data collected by the Department of Finance, was equally ambiguous. It indicated that while evaluation of some type was being referred to in the preparation of both Finance and portfolio budgetary advice, it was not on the systematic basis envisaged by the strategy; that is, it was referred to on an *ad hoc* basis when it was available. Further, there was no evidence to suggest that this evaluation was *more* oriented towards how effective programs had been in achieving their objectives. One could surmise that part of the problem lies in the disjointed nature of the evaluation cycle *vis-a-vis* the budget cycle combined with predictable resistance tactics on the part of a number of portfolios. At the same time, the political executive, the ultimate user of program evaluations, also showed signs of dissatisfaction. As discussed earlier, evaluation was part of an exercise in expenditure restraint, but these objectives, particularly after accumulated experience, were not served as anticipated. As a result, sharper instruments—in the form of program reviews—were introduced.

The above reservations hold a number of implications for both the future of the evaluation strategy as a budgetary tool as well as the resource coordination dimension of central agency policy management. First and foremost, the above analysis suggests that, as the central budget agency, the Department of Finance has to decide what its core business is in relation to resource coordination. Accepting that policy and expenditure are two sides of the same coin, is the central budget agency better off staying out of the analysis of program effectiveness and hence limiting its contribution to the assessment of cost-efficiency, as one recent study has suggested (Uhr 1996 (forthcoming), pp16-21)? The evidence contained in this chapter suggests that the role of supply officers as policy analysts is a traditional and enduring one. Precisely because of the relationship between policy and expenditure, the role of the budget agency will remain locked into the review of policy on its *merits* rather than adopting a purely financial perspective. The informality of the budgetary process itself dictates the degree of success which a system establishing a purportedly formal information stream, such as the evaluation strategy, can attain. These issues are explored further in the concluding chapter (Chapter Eight).

The second task facing the Department of Finance is to satisfy itself that the evaluation strategy is serving these resource coordination purposes properly. In terms of its budgetary objectives the evaluation strategy has been severely handicapped by placing itself in a centralisation-decentralisation bind. In this regard the Auditor-General’s critique of the Department of Finance’s handling of evaluation planning holds some validity. If Finance is committed to integrating program evaluation with the budgetary process it should give consideration to some of the following issues.

First, central budgetary decision making should be promulgated as an *explicit* policy. While use in the budgetary context was seen by *Finance* as the major objective (see Chapter Five), this was continually played down for fear of undermining portfolio compliance. At present the evaluation strategy is a portfolio centred policy. This is perfectly acceptable if evaluation is to serve portfolio (primarily program management) purposes, but under these circumstances, as one supply officer pointed out, Finance has to accept that it is basically a "parasite on the back of evaluation" (Interview 06). Second, either the evaluation planning process must be centralised so that Finance is better able to determine the evaluation agenda, or the devolved evaluation planning process should be tightened by attaching real penalties to non-compliance. In other words, the central budget agency must synchronise the evaluation planning and budget cycles. Third, and as a corollary to the first two points, program evaluation must be tailored to the information needs of the user, in this case supply officers (as intermediate users) and ministers (as end users). The Australian evaluation strategy was admirably comprehensive in its scope, but this comes at the cost of spreading the benefits of evaluation too thinly across the administration and ultimately to the detriment of the *competing* purposes. If evaluation is intended to play a part in central priority setting there must be a recognition that this is essentially a coordination role and that the responsible central agency must ultimately play an enforcement role. The problem, of course, is to determine the range of this enforcement—under what conditions does it make sense for line departments to trade information with central agencies?

Finally, the budgetary use of program evaluation highlights how important the *informal* aspect of review is in the work of central budget agencies. It was noted earlier in the chapter that evaluation was an attempt to formalise aspects of the analytical review work of supply divisions. A crucial problem appears to be that the formal aspects are heavily dependent on the informal processes, which comprise the judgment, intuition and experience of supply officers in identifying emerging policy priorities, suspect programs that may require investigation, possible targets for savings options as well as the accompanying negotiation skills. The way in which the evaluation planning game relied on the bureaucratic cunning of both supply officers and portfolio officers provides one illustration. On reflection the evaluation strategy could be interpreted, at the budgetary level at least, as having laid its foundations in sand—it was based in a weak capacity to formalise processes that are not, due to the nature of the work, amenable to formalisation. In this way, given the questionable impact of the evaluation strategy on aspects of the budget process, efforts to augment resource coordination in the Department of Finance have seemingly been running to stand still.

Chapter Seven

Returning To Policy Process: Policy Coordination And The Evaluation Of Policy Advice

The focus of this chapter is on the recent development of the evaluation of policy advice as both a distinct component of the evaluation strategy, and perhaps the most instructive instance of evaluation serving the policy management functions of central agencies. The main argument is that the evaluation mechanism adopted, the policy management review (PMR), has emerged as a tool for increasing the influence that central agencies have in the *processes* that generate policy advice. In stark contrast to the kind of “standard” program evaluation discussed so far, the evaluation of policy advice was an area where the rhetorical concentration on the outcomes of programs was diluted. Instead, the central agencies were able successfully to steer evaluation away from questions of public accountability for the outcomes of policy advice, towards arrangements for achieving more effective control of the processes that underpin the production of advice.

Increased control, however, was not the underlying motivation for evaluating policy advice. The PMR format represented an attempt by central agencies to preserve the credibility of the evaluation strategy within both the public service and the wider community. The return to process was a product of the unresolved tension between public service and ministerial responsibilities for policy that the evaluation strategy had inherited. In effect, extending the evaluation strategy to policy advising programs marked out the limits to evaluating the outcomes of government activities. In arguing this, the chapter proceeds in two steps.

First, the impetus for the evaluation of policy advice is examined as well as the difficulties facing its development, particularly in terms of determining what to evaluate and how, and who actually was to do it. Because policy advice was equated specifically with central agencies, the nature of the policy advising task in these agencies is briefly described as well as the claims each has to distinctiveness in this area. As we will see, the PMR format was a compromise position that sought to placate both demands for public accountability and, in some cases, intense central agency opposition to opening up their policy advising work to outside scrutiny. Second, it analyses the conduct of PMRs and speculates on their potential use for central agency policy coordination. It can be recalled from Chapter Three's consideration of the literature on coordination that "policy coordination" could be equated with the management of conflict in policy making. PMRs have consistently been framed around the interdepartmental committee (IDC) process, where policy development and hence coordination hinge on the ability of participants to achieve compromise between competing interests. The most likely role for PMRs may be to give central agencies more leverage in managing these processes of compromise.

THE POLITICS OF EVALUATING POLICY ADVICE

Until the early 1990s the policy advising programs within both central agencies and line departments were the exception to the program evaluation rule. The obligation imposed by the evaluation strategy periodically to evaluate all programs had full application to those agencies whose chief task was advising government, but throughout the term of the FMIP the measurement of performance in policy advising, particularly effectiveness, had been, like the PMB initiatives before it, consistently met with scepticism in most areas of the bureaucracy. Certainly, as early as 1986, the Department of Finance had countenanced the assessment of performance in policy areas as an integral part of the development of PMB—policy areas were to be identified as "responsibility centres" within program structures and hence obliged to manage themselves according to the dictates of the management cycle (which, of course, gave prominence to evaluation) (DoF 1986b, pp1, 5). However, even at this early stage, policy programs were to be insulated from the full rigours of evaluation, on the basis that performance assessment could only be achieved through the development of work plans against which progress might be monitored in terms of cost, quality, timeliness and relevance "as judged by peers and Ministers" (DoF 1986b, p6). Effectiveness was discounted as an "inadequate" measure of performance in policy advising.

This ambivalence—a consistency in treating policy advice like any other program coupled with an inconsistency in granting immunity from effectiveness evaluation—was to fetter the application of evaluation to this area for the remainder of the decade. Throughout this period, the affected departments commonly protested that evaluating their policy advising activities was impractical since it presented what they saw as insurmountable problems of definition (that is, what to evaluate against which criteria?) and fuelled concerns that, if pressed, assessment would inevitably fall back on "judgemental factors" (PSB & DoF 1986, pp33-35; DoF 1988b, pp130, 146, 160). In any event, the evaluation of policy advice was not pressed, although not solely because of methodological obstacles. In line with the type of budgetary concerns that so comprehensively dominated program evaluation, the Department of Finance instinctively framed the evaluation strategy around those programs that held the "largest resource implications" (ANAO 1991a, para 3.2.10, p33). In budgetary terms, policy programs were minnows that, in the absence of external prodding, could well have been destined to remain outside the evaluation strategy. As this section will show, efforts by central agencies to evaluate policy advice can be seen as an anticipatory response to parliamentary "threats" of intensified scrutiny. At the same time it was a response that was shaped by and for central agency coordination interests.

It is perhaps not all too remarkable that, given the expenditure policy interests of the Department of Finance, the evaluation of policy advice was eventually promoted by the Department of the Prime Minister and Cabinet (DPMC) (although this had as much to do with senior personnel changes as departmental perspective). As discussed in Chapter Five, a fundamental reconsideration of the evaluation strategy at the beginning of the 1990s was prompted by the stinging criticism contained in the first stage efficiency audit of evaluation conducted by the Audit Office. The evaluation of both policy advising and regulatory programs had been singled out as "areas of neglect" and the report called on the Department of Finance, as steward for the strategy, to maintain policy consistency by developing evaluation procedures specific to the policy advising function (ANAO 1991a, paras 3.2.8-3.2.11, pp32-33). Given the lack of progress in policy evaluation, each of the agencies subjected to the audit faced little choice other than to subscribe to this course of action, although

DPMC's response was somewhat circumspect, cautioning that the risk of devoting substantial resources to such a difficult area of evaluation might only produce limited benefits (ANAO 1991a, para 3.2.13, p34; Codd to Taylor, February 1991. FD: 91/1828).

Taking the lead provided by the Audit report, other parliamentary fora questioned the adequacy of performance reporting by central agencies on policy advising programs. Senate Estimates Committees, for instance, were independently voicing two types of dissatisfaction with the way in which the central agencies, and in particular DPMC and the Department of the Treasury, were approaching their public accountability obligations under the evaluation strategy. The first was directed at the form in which policy program performance was reported. This was judged to be

> . . . dreadful. They are waffly. They are full of rhetoric and jargon and it is very hard to get beyond that to find any real substance (CPD, Senate Estimates Committee A, 25 September 1991, ppA464; see also 9 September 1991, A127-128).

The second recognised that form followed function. From portfolio performance statements, and in particular those of DPMC, the most common way of evaluating policy advising programs was the confidential 'ministerial-satisfaction' test. In Senate Estimates hearings it was argued that, as an assessment mechanism, ministerial satisfaction was patently insufficient in terms of public accountability, since there was no objective standard against which the public could assess the quality of advice (CPD, Senate Estimates Committee A, 9 September 1991, pA130). In these cases there would effectively be nothing to report except that the minister was pleased with the department's performance. In other cases, the minimum formality (and maximum self-assurance) of the assessment procedures *presumed* that advice could be nothing other than of the highest quality.[1] In either case the evaluation of policy

[1] Informal assessment of this type was commonly reported in Senate Estimates hearings. The following exchange, involving a senior officer from the Treasury Department, is indicative:

> **Senator Short—**On page 71 [of the budget estimates] you mention the facilitation of worthwhile foreign investment in Australia [as one objective of your policy advising programs]. How do you define worthwhile foreign investment?

advising programs was, as an issue of public accountability, emerging from the shadows of the evaluation strategy.

A Common Distinction? Policy Advising In The CentralAgencies

Before we examine the momentum for, and outcomes of, efforts to extend the evaluation strategy to policy advising programs, we must first gain some understanding of the nature of policy work in the three primary central agencies in Australian government—DPMC, the Department of the Treasury, and the Department of Finance. Of these, the policy work of the first two only will be considered here, since the policy advising responsibilities of the Department of Finance were covered in the previous chapter. The aim is to demonstrate how each of the central agencies contend that it has a distinctive policy coordination role which determines the complexion of its policy work. This influences how each agency views both the policy advising process and the nature of the advice that it ultimately generates. More importantly, this adherence to distinctiveness goes a long way in explaining much of the bureaucratic opposition that confronted policy evaluation.

All three central agencies concede that there are aspects of the policy advising task that are common. There is general recognition that, by virtue of their location at the apex of government decision making, central agency policy advising responsibilities can be categorised into two types (DPMC 1989; FD: 91/1828). The first of these, *general* policy advising, is characterised by detailed commentary on proposals initiated by other departments (for example, those

---

Mr Pooley—That is only a reference to the fact that one or two proposals a year are rejected by the Government as contrary to the national interest, but the others that are approved are encompassed in the word “worthwhile”.

Senator Short—So, by definition, “worthwhile” is anything that is not rejected; is that what you are saying?

Mr Pooley—Yes. Pretty close.

Senator Short—Given that there are virtually no rejections you are saying therefore that—

Mr Pooley—Most of it is worthwhile (CPD, Senate Estimates Committee A, 25 September 1991, pA546).

submitted for budget Cabinet). It involves applying a "whole-of-government" perspective that assesses the proposal against prevailing government priorities and apprehends the likely impacts on other departments and policy sectors. This task is also referred to as *reactive* policy advising, since it often emerges as a response to changed policy conditions or the altered political demands of the government (Interview: 08, 21). The second category is *specific* policy advising. In this situation the central agency is required to generate its *own* policy initiatives. This task is similar to policy advising in line portfolios except that there is no attached implementation component. The balance between the two advising tasks is determined by the role each central agency assumes.

At the most general level, the primary tasks of DPMC comprise the coordination of government administration and advising the prime minister across the full range of federal responsibilities. In carrying out these tasks, DPMC's approach will always reflect both the political and managerial style of the prime minister, and the way he or she conducts Cabinet:

> The DPMC serves the prime minister's needs as an individual leader and the leader of a team. It takes the lead in certain policy areas of the prime minister's choosing and maintains a watching, coordinating brief on all others. This can be done with a light or heavy hand, according to the personal style of the prime minister. . . (Warhurst 1992, p511).

However, DPMC responsibilities in policy making, regardless of prime ministerial style, are distinct. According to one recent secretary, these responsibilities are twofold (Keating 1993b, pp6-9; 1995, p4). DPMC is, first and foremost, a custodian for *due process* in policy advising. In this capacity it must guarantee "the development and application of processes designed to ensure that all legitimate interests are considered and decisions properly informed". In addition to securing procedural integrity, it also takes a leading role in building *contestability* into the government advisory system, by ensuring that an alternative whole-of-government approach is reflected in policy development.

This dual advisory role requires DPMC to "add value, not to duplicate" (Keating 1993b, p9). So how does DPMC purport to "add value" to the policy advising process? Policy coordination is most often called for when particular policies have an impact beyond their own sphere, crossing departmental divisions. It is for this reason that Cabinet is likely to involve itself in consideration of the policy and/or defer the issue for consideration by an IDC that is led by one of the central agencies, commonly DPMC. What is distinctive about DPMC's contribution is that in coordinating policy development it continually tests the merits of departmental advice from a perspective other than the financial or economic. This is what Keating terms the "broader policy agenda":

> The Budget process is a traditional and powerful tool for [ensuring the overall coherence of Government policy]. There are, however, some risks for policy coherence if policy making is totally dominated by the Budget process . . . One of our challenges is to achieve a balance between the broader policy agenda and budgetary imperatives, and to ensure that the various decision making processes facilitate that end (Keating 1995, p3).

Even though this refers specifically to the budget process, the argument applies to all DPMC advising work, particularly the IDC process. It attributes a distinctiveness to the DPMC policy advisory function that places a premium on the contestability of policy advice generated either within or outside government, as well as playing a broker role in seeking to balance social and economic perspectives in policy development.

Although some consider that the Treasury Department no longer commands the title of central coordinating agency—the Department of Finance is now the "principal second voice in government" (Interview: 27)—there is no doubt it continues to exercise considerable influence over policy coordination. This stems from the Treasury's advising responsibilities on a number of complementary policy areas, including macroeconomic (fiscal, monetary and wages) policy instruments, as well as microeconomic or structural polices, such as trade and industry assistance, regulation of financial markets and taxation policy. The contribution that the Treasury makes to policy development and coordination is pervasive; its advice frames most if not all government policy options. In this way the Treasury argues that its policy advising role can be distinguished because it takes in such a "broad sweep" (Higgins & Borthwick 1988, p61; cf Treasury 1991, p6). This posture has been described by one former secretary in the following terms:

> Treasury advice is very broad in its perspective. More than other departments, the Treasury endeavours to take an economy-wide view and to be concerned with "the general interest". The Treasury does not have a narrow "clientele" in the sense that most other departments do. It should not have particular sectional barrows to push . . . It is this responsibility for looking at the economy as a whole—"the general interest"—which mainly distinguishes the Treasury from other departments (Fraser 1984, p231).

There is then a distinctiveness to policy advising functions that appears to be common to the central agencies—each purports to take a "broad view" of policy that accords with both government priorities and "the general interest". In the task of bringing coherence to government policy there must be some truth in this argument, although it is perhaps more accurate to say that each central agency brings a *distinct piece* to the coordination puzzle, without which the policy advising process would not be complete or properly balanced. These constitute the various dimensions of the policy management function, *viz* resource coordination, and policy development and coordination. As the next section will show, however, adherence to the "common distinctiveness" argument was a chief factor in slowing down the application of evaluation strategy to policy advising programs.

The Evaluation Crucible: The Working Group On Performance Assessment Of Policy Work

Despite a number of misgivings that it had placed on the public record, as well as Finance's clear functional responsibilities for program evaluation, DPMC moved quickly to fill the policy void around the evaluation of policy advice. In April 1991 the then Secretary to DPMC, Michael Codd, initiated a meeting of representatives from DPMC, Finance, Treasury and the Public Service Commission (PSC) directed at developing measures for assessing policy work within central agencies. These would be measures "that would not require excessive and artificial recording, but which would be seen as meaningful by Parliament and external watchers" (Codd to Keating, April 1991. FD: 91/1828). It was from this meeting that a proposal for a high level Working Group on the Performance Assessment of Policy Work emerged (hereafter the Working Group).

The Working Group, which was established during May and June of 1991, was required, by the close of 1991, to "examine and report on ways of assessing the performance of policy work which would be credible for the purposes of external scrutiny". These terms of reference certainly confirm the views of one senior member of the Working Group that its establishment was generated by "parliamentary pressure, particularly on central agencies, to demonstrate that they were adding value to the process of government" and hence constituted a "response to a threat" (although the nature of this threat was not so clear) (Waller 1996, p69).[2] At the same time, the Working Group was also dominated by DPMC. This is illustrated by the working group's composition—not only was it chaired by senior DPMC officers, first Mike Lawless and then Mike Waller, but four of the seven positions went to DPMC officers. In part this level of participation reflected the interests of the new Secretary of the Prime Minister's Department, Dr Michael Keating (Interviews: 08, 12, 21); it also reflected concerns on the part of DPMC to ensure that the framework for policy evaluation was more sympathetic to policy coordination tasks than the narrow resource coordination tasks stressed by Finance. It was no accident that on the development of policy advice evaluation "Finance was left behind" (Interview: 10).

As the central objective of the Working Group was to develop an evaluation framework that could demonstrate the *public* credibility of central agency policy advising programs, DPMC produced a paper designed to focus discussion on key public accountability issues. The paper identified two crucial concerns with extending the evaluation strategy to policy advising programs—first, determining to whom public service advisers should be held accountable, and second, what they should be held accountable for, and how (DPMC 1991, p2). Noting that the links between policy advice and the outcomes of policy were "at best indirect", DPMC suggested that two alternative approaches to evaluation presented themselves (DPMC 1991, p5). The first sought to evaluate the policy advising *process*, which appeared "more feasible and appropriate", especially where the link between the advice and its outcomes were unclear.

[2] There has also been some suggestion that the evaluation of policy advice represented not just an attempt to improve policy by demonstrating the value of the public service to the community and Parliament, but that it was also part of a *defensive mechanism* designed to ensure the "survival" of the public service in the face of diminishing political and community support (Record of Conference Discussion in Uhr & Mackay 1996, pp130-131).

The second concentrated on evaluating the *outcomes* of policy advising. This was considered more "problematic" than the first route since it necessarily required evaluating "discrete impacts" that might conceivably include not only the effect on the recipient and the extent to which the advice was adopted, but, if adopted, also whether the final result could be attributed to "specific agents" (DPMC 1991, pp7-8).

This methodological distinction between the processes and outcomes of policy advising dominated the first report of the Working Group, which was completed in December 1991.[3] The primary reason why the distinction held sway can be found in the constraints imposed by the model of "Westminster cabinet government" which the report took as its reference point (PSC 1995, p1). Even though the report recognised that there were effectively two impulses for evaluating policy advice—the establishment of a system for holding policy advisers publicly accountable, as well as the potential for improving the quality of policy work—fixing the limits of public accountability for public service advisers would ultimately determine the scope for improved policy development and coordination.

In its report, the Working Group refines the distinctions outlined by DPMC in its issues paper by charting the policy advising process and identifying those parts that may be more amenable to assessment. Three components were distinguished at the product end—outputs, policy and outcomes. Output, such as a brief prepared by policy advisers, was marked off from both policy, which was defined as "a course of action decided by Cabinet or a minister intended to accomplish some end", and outcome, or the "the real world impact of the policies in terms of what happens in society" (PSC 1995, pp4-6). The main problem in evaluating policy advice was deciding the point at which effectiveness could best be judged. Should it be when the recipient either accepted or rejected the output and decided on policy, or at a time when the outcome of the policy can be determined? In line with the conventional Westminster model that figured so prominently in contemporaneous MAB-MIAC pronouncements on public accountability (MAB-MIAC 1991, pp2-4, 10;

[3] The first report of the Working Group did not receive a full public airing until September 1995 when it was published as part of the Public Service Commission occasional paper series (PSC 1995). An abridged version of the report was, however, published in 1992 under the name of the second convenor of the Working Group, Mike Waller (Waller 1992).

1993c, pp6-7, 13-14), the report safely locates ownership of, and public accountability for, "the main outcomes of policy" with ministers, and stresses that "assessment mechanisms have to be designed which protect the confidential relationship between ministers and their advisers" (PSC 1995, pp11-13). In this way the report retreats from a primary focus on the outcomes of advice to the processes of advising, by asserting that the desirable characteristics of policy advice outputs can *only* be identified "by reference to the nature of the policy formulation process"—in order not to violate "confidential policy dialogue" a system of external evaluation must illuminate the "quality of the policy advising process" (PSC 1995, pp10, 14).

Advocating a "cautious and experimental approach to innovation", the report canvassed three options for a policy advice evaluation system (PSC 1995, pp24-26). The first revolved around assessments *internal* to government and designed to supplement existing informal mechanisms that seek to enhance policy advising capacities. The second constituted a more formal version of the first in that a peer review would report confidentially on the management of the advising process. These first two options, although directed at improving the conduct of central agency policy work, clearly failed to meet public accountability requirements. The third option was to develop a compromise methodology which the report termed policy management reviews (PMRs). It was envisaged that PMRs would engage assessors external to, but experienced in the work of, the bureaucracy, and follow a formal assessment procedure consisting of both quantitative and qualitative methods that

> would look at the environment of a policy area to examine its management and ascertain whether it has characteristics conducive to a good policy process. Utilising particular case studies, the reviewers would focus on the policy advising process, the outputs of the process and the ultimate outcomes—where they can be linked to the policy advice given (PSC 1995, p26).

Although PMRs were shaped as assessment mechanisms for process, outputs and outcomes, in the context of the report's rejection of public service responsibility for policy and outcomes, the proviso for discernible linkages between the three all but excludes evaluation of these end products. This is reinforced by the proposal for the PMR format to depart from the evaluation strategy's requirement for continuous evaluation. A projected scarcity of qualified assessors and the "need for selectivity" in choosing examinable policy issues ruled out systematic evaluation (PSC 1995, pp27-28). It also conceded that process was the common (and hence politically feasible) element of policy work that was open to assessment.

The Working Group's report remained in draft form until April 1992, when it was considered and finally approved by the secretaries of the four participating central agencies (Gregory 1996, p141). In June and July, the secretary of DPMC submitted a final version of the report to Prime Minister Paul Keating and, in conjunction with the secretary to the Department of Finance, secured the agreement of both Keating and the Minister for Finance, Ralph Willis, to undertake a series of trial PMRs within central agencies (FD: 91/1828). The terms of reference for these trials, although adhering in form to the "efficiency and effectiveness of the provision of policy advice", were, in substance, narrowed to the adequacy of both the management of resources and staff, as well as processes of consultation with affected community organisations, departments and ministerial offices (FD: 91/3908/2).

Although DPMC had dominated the process of policy formulation in the area of policy assessment, the lapse between completion and endorsement of the report reveals much about the level of discord operating within the Working Group. According to the second convenor of the Working Group, the report was left in draft form for two reasons (Convenor of Working Group Mike Waller to Secretaries of Central Agencies, December 1991. FD: 91/3908/2). There was, of course, a natural element of caution on the part of the Working Group to dampen the rate of change, since "once the recommended course has been embarked upon there will be difficulty in withdrawing". This approach was, however, chastened by the "divided views" of the Working Group and, in particular, the "significant reservations" that the Treasury held about the paper's tentative conclusions. In retrospect, as the next section will show, Treasury's discomfort with the report now appears somewhat affected; by and large, the report's conclusions substantially reflect the Treasury's core objections to evaluating the outcomes of policy advice.

The Search For Common Ground

The final report of the Working Group may have given the appearance of settling on process as the most appropriate focus for policy advice evaluation, but this position was achieved only after a protracted gestation during the life of the Working Group, and continuing disagreement afterwards. To some degree each of the central agencies remained sceptical about their own capacities to evaluate policy work in a credible (that is, public) fashion and, perhaps more significantly, what such assessment could actually hope to achieve. To some extent this is an indication of the rivalry among the departments and the tension produced by the potential for evaluation to undercut the *raison d'etre* of these departments. The most intense—and public—opposition came from the Treasury Department, with one senior Treasury officer conceding that the department “did not make a good contribution” to the Working Group process (Interview: 21). There were, however, a number of issues that served to splinter the central agency consensus.

Immunity Based On Policy Distinctiveness

It was suggested earlier that the definitive commonality of policy advisory work in the central agencies could be found in the so called distinctiveness of the policy management tasks that each of the departments believed it faced. In extending evaluation to policy programs, this same position was frequently invoked, to good effect, as a primary reason *against* evaluating the effectiveness of policy advice. Even DPMC, which, according to observers in other central departments had taken a “monumental leap” in initially pushing for the evaluation of policy outcomes (Interview: 10), rationalised the return to process in terms of its whole-of-government policy responsibilities. For instance, the secretary to DPMC took the position that the value of evaluating policy advice lay firmly in his department’s particular obligations with respect to policy coordination:

> the whole-of-government perspective of DPMC is distinctive. Finance and Treasury claim they come close, but expenditure reduction is an important although not the only objective. Treasury represents the private sector in Cabinet, Finance represents the public sector. Some issues do not have economic imperatives (Interview: 12; cf Interview 08; compare with quotation on p186 above).

As a consequence, "DPMC should be a strong protagonist for good process [rather than being] a protagonist for policy positions", and for this reason makes no apologies for a PMR format that concentrates on process rather than policy outcomes (Interview: 12; Record of Conference Discussion in Uhr & Mackay 1996, p115).

Other central agencies, however, were not as comfortable with opening up their policy advising processes. The Department of the Treasury, while acknowledging that in principle policy evaluation was meritorious, held the position that PMRs could not be done for the "80 per cent" of Treasury policy work which was "quick and dirty". This was the view of senior Treasury officers who contend that the economic policy advice offered within the Treasury is highly confidential, subject to rigid time constraints and quite often provided within a few minutes either orally or "on the back of an envelope" (Interview: 01, 21, 27).[4] As a consequence, oral requests from ministers or ministerial staff do not have separate files and may contain policy briefings which the government may not want public comment on. The implications for assessment of policy work are obvious. How can policy work be evaluated if there is no paper trail and, according to Treasury officers, public accountability bodies harbour a "fixation on the paper trail for policy development" (Interview: 21, 27)?

This component of Treasury opposition was played out during a project audit of the Treasury's procedures for managing its Economic Policy Program (ANAO 1992b). The Audit report, which flatly criticised the Treasury for its failure to develop adequate performance indicators to measure the quality of its policy advice, also confirmed the informal methods of its policy advising. The servicing of ministerial requests for economic policy advice and research work was characterised by the absence of either recorded details of planned performance or summary data of completed work (ANAO 1992b, paras 4.8-4.9, 4.11-4.12, p4). Policy processes of this type did not serve either internal policy management or external accountability purposes. The Treasury response was to deflect the burden of proof onto external observers. It would be "impractical"

[4] The mode of advice delivery—that is, either oral or written—is highly dependent on ministerial style. This would, according to Treasury officers, influence how any assessment could be undertaken. As an example, they cite Paul Keating's time as Treasurer and that Minister's preference for informal, oral based advising (Interview: 21).

for the Treasury to formalise a work practice of recording requests for advice "given the very large number of day-to-day requests it handles". Instead, external scrutineers should provide "clear evidence that current arrangements are not working satisfactorily" (ANAO 1992b, para 6.7, p6; cf Treasury 1992, pp64-67).

Treasury's position on the integrity of its own informal practices of policy advising and evaluation was also adopted by other central agencies. The Department of Finance, at least initially, pointed to its own internal system of formal and informal assessment of expenditure policy advising which already relied on peer review mechanisms (FD: 91/1828). These included the circulation of policy briefs for peer and executive consideration, meetings between departmental officers and ministerial advisers and Post-Budget seminars which analysed performance in budget-related activities. Similarly, the Department of Foreign Affairs and Trade (DFAT), which is composed largely of policy advising programs, maintained that formal evaluation would not displace the significance of internal performance reviews of overseas posts and Canberra-based divisions (see Forrester 1993, pp161-162; DFAT 1991, pp52-53). According to one senior officer in DFAT, the role of that department is "distinctive" in that much of the policy work undertaken by its divisions is indivisible—formal evaluation would "basically be process for its own sake" since the kind of review envisaged by PMRs is effectively already being done (Interview: 09). Forced redundancy in the review of policy work loomed large as a continuing departmental objection to PMRs (although this ignored pressures for increased public accountability).

What To Evaluate And How

The evaluation of policy advice presented a number of challenges for the evaluation strategy, not least of which was deriving, essentially from first principles, a methodology that was acceptable to all the central agencies. Earlier, special mention was made of the Treasury's objections to evaluating the outcomes of policy advice. This defensive posture was stirred, in part, by the widespread media criticism of Treasury's advice on monetary policy in late 1989 and early 1990 (see Treasury 1992, pp63-64). In this case the Treasury had been charged with both seriously misjudging the balance of payments constraint and inexcusable tardiness in advising the Treasurer on the reduction of interest rate levels. The unsatisfactory policy outcome, which commentators attributed directly to the Treasury advice, was the prolonged economic recession of the early 1990s (Kelly 1994, pp487-489, 495-496).

It was against this backdrop that the Treasury, somewhat brazenly, publicly voiced its reservations about extending the evaluation strategy into the realm of policy work (Treasury 1992, pp62-67; more recently see Treasury 1995, pp16-17; MAB-MIAC 1993a, p377). Following closely the Working Group's distinction between policy work outputs and outcomes, the core of Treasury's opposition was its insistence that there could be no "practical and objective means" for evaluating the effectiveness of policy advice, since assessing the merit of advice outputs was "obviously a matter of judgment which clearly will vary from one individual to another" (Treasury 1992, p63, 101; Treasury 1993, p104). This judgmental factor weighed heavily:

> *ex post* evaluation is, by definition, made with the benefit of hindsight. In sharp contrast, policy advice is framed within the context of circumstances prevailing at the time. In theory it could be argued that *ex post* performance evaluation should properly assess the quality of policy advice according to the then prevailing circumstances, but *eliminating the influence of more recent information will always remain problematic* (Treasury 1992, p63, emphasis added).

For the Treasury, the assessment of policy program effectiveness was beyond the evaluation strategy precisely because the link between policy advice outputs and outcomes is, at best, indirect and, more significantly, affected by a host of intervening factors that make the attribution of policy "effects" too onerous (similar views were expressed by DFAT and DPMC, see FD: 90/3210). These factors may include unrelated policies that have an unintended (if not unforeseeable) impact on the policy in question, as well as the prickly task of determining the appropriate "time frame" for the evaluation—that is, at what point after the policy decision is made should evaluation occur in order to measure the "full implications" of the policy?

Arguments of this type, which flatly denied the possibility of methods for evaluating the end products of policy advice, continued long after experimental PMRs had been initiated in central agencies (see, for instance, MAB-MIAC

1993a, p377; Treasury 1995, pp16-17). In the final analysis, this articulate and public opposition ensured that the PMR format was the product of compromise. The omission of policy programs from the evaluation strategy had slowly given rise to a perception held within both the central agencies and the public service more generally that the central agencies were failing to practice what they preached. Certainly, some sections in the Department of Finance were aware that evaluating policy programs might lend more credibility to the strategy in the eyes of line portfolios (Interview: 15), although others feared that PMRs raised the prospect of "more potential bureaucratic monsters" (Finance Division Heads Meeting. FD: 91/3908/2).[5] At the same time, the source of any growth in resentment could just as likely be traced to other central agencies, such as the Treasury and DPMC, which accused Finance of failing to "get its act together" on policy evaluation (PMR Roundtable, October 1993. FD: 91/3908/2; Interview: 21).

Public Accountability And The Attribution Of Responsibility

The final area of contention was what, from the perspective of central agencies, constituted the optimum level of public accountability for the evaluation of policy advice? If the intended PMR format was indeed an extension of the evaluation strategy, then the strategy's presumption in favour of public release (which also had the support of a Cabinet decision) must also apply. Automatic coverage, however, was not assumed, there being considerable angst over the trade-off between the benefits of public accountability and the costs in terms of the utility of the reviews to departments. Again the Treasury, this time joined by DPMC, argued that at the early "foetal" stage, the reviews should not be published (PMR Roundtable, August 1994. FD: 91/3908/2). Objectivity (or, more accurately, candour) in the review process was seen as the higher priority, and the only way to ensure this was for reviewers to consider their work as an assessment task internal to government management rather than an evaluation

[5] It was not clear whether this last warning related to the prospect of opening an unstoppable drain on administrative resources (although this is certainly hinted at when some Heads commented that the Working Group paper had the "flavour of a treadmill of evaluations"), or the fact that PMRs would unnecessarily open a new front on political sensitivities not specifically sanctioned in the evaluation strategy (FD: 91/3908/2).

task directed at satisfying, no matter how meagrely, accountability obligations. The two departments concurred that "'Accountable to whom?' is not yet a question we've resolved to the point at which agencies approach evaluations with a genuinely inquiring mind" (PMR Roundtable, August 1994. FD: 91/3908/2).

Again the least receptive of the central agencies proved to be the Treasury, which was pushing strongly for PMRs as a tool for internal management purposes. It regarded any obligation for public release as self-defeating, since it would ensure a "bias in the choice of PMRs" towards those programs and policy areas less susceptible to criticism (Interviews: 21, 27). The end result would be negative on both accountability as well as management improvement counts. As the next section will show, the Treasury applied this reasoning in maintaining the confidentiality of its first PMR.

But the principle of public accountability was only the starting point for Treasury resistance; the primary concern was the problem of attributing the different components of policy advice to specific advisory agents, something which the Treasury believed the report of the Working Group had not adequately addressed. More specifically, even if PMRs were to focus on the process of policy advising, the Treasury held grave misgivings about the capacity of reviewers to attribute the *inputs*, let alone the outputs, of policy advice. In short, there were serious drawbacks with attempts to evaluate process, in that they were also judged to have both misread the nature of the advising process in contemporary Australian government, and failed to accurately identify *where* policy advice influence is now exercised. For senior Treasury officers efforts to evaluate public service policy advising would be self-defeating unless the role of ministerial advisers and their relations with the administration were clarified (Interview: 21, 27).

There appear, therefore, to be questions about the extent to which policy advice input can be attributed to different contributors, and that these doubts arise most acutely in the interface between departmental advisers and those advisers located in ministerial offices. The latter are perceived by public servants to exercise considerable influence over what advice ministers actually see and hear, but are subject to few, if any, formal public accountability obligations. In this way the role of ministerial advisory staffs are pivotal since they often shape the advice that is generated by departments. As an example, Treasury officials cite the forecasts contained in annual Budget papers which are prepared by the Treasury in the name of the government, but which are, in policy substance, written by economic advisers located within the relevant ministers' offices (Interview: 21, 27).[6] The most recent study of relations between ministerial advisers and departments certainly confirms the potential for this influence (Dunn 1995). While advisers and departmental secretaries contend that independent policy development work is undertaken by departments and not ministerial offices, the roles played by the two agents are complementary, since each brings different strengths to the policy advisory process—advisers are more aware of the political implications of specific courses of action, while public service advisers are more familiar with particular policy issues and the problems of implementation. Nonetheless, ministerial offices were often seen to "filter" advice sourced from departments, packaging it so as to minimise the potential political damage, and hence influencing the advice ultimately received by ministers (Dunn 1995, pp509-512; see also Walter 1986, pp146-150, 152-157).

For the purposes of public accountability, the problem of attribution, and, as an instructive example, the blurred role of ministerial advisers in the development of policy advice, served not only to reinforce the hazards posed by evaluating policy advice outcomes, but also signalled the restrictions that might be met in assessing policy advisory processes. Nonetheless, evaluating process was the compromise position arrived at by the central agencies in the report of the Working Group. It sought to accommodate the competing pressures associated with early (and, in the case of the Treasury Department, continuing) 'in-principle' opposition to examining the outcomes of policy decisions, the absence of an acceptable methodology for assessing policy programs, and a shift in policy design towards policy advice evaluation as a tool for assisting policy development and coordination and away from the initial concern with public accountability issues. In the end the PMR format was an attempt to satisfy these public accountability obligations with a review process that

---

[6] This assertion received further support at the time of interviews when a former secretary to the Treasury Department, Tony Cole, claimed that the highly contentious income tax cuts announced as part of the Labor Government's 1992 *One Nation* economic statement were devised *without* the detailed advice of the Treasury (*The Canberra Times*. 9 May, 1995. pp1-2). Cole maintained that the Government's economic forecasts in 1992 were "predominantly influenced by people in the Prime Minister's office" rather than the Treasury Department, and that the forecasts themselves had, as a consequence, "strained credibility".

possessed, first and foremost, internal utility. With this in mind, the purpose of the next section is to consider the results of the first PMRs and assess their potential for enhancing central agency policy development and coordination through analysis of policy process.

POLICY MANAGEMENT REVIEWS: RETURNING TO POLICY PROCESS

Following the Government's approval for review trials in mid 1992, three of the central agencies represented on the Working Group commenced a series of experimental PMRs. These are listed in **Table 7.1** below. Of the five PMRs completed between 1993 and 1995, four were made available for public release. At this point it is noteworthy that the Treasury's PMR retained a confidential status, but more on this later. As a consequence of these restrictions, this analysis is forced to concentrate on the four PMRs that are publicly available. Again, from **Table 7.1**, we can see that three of the four reviews take as their focus interdepartmental committee (IDC) policy processes, and this fact is used to frame discussion in this section, the aim of which is twofold. First, it examines why PMRs have been directed at IDC processes and briefly assesses their fitness for this task. Second, it identifies the common elements of PMRs and analyses their potential for increasing the effectiveness of central agency policy management functions.

Table 7.1 Central Agency Policy Management Reviews 1993-1995

	Review	**Object of Review**	**Department**	**Focus**
	Glenn Review April 1993	1992-93 Carers Package	DPMC	IDC policy processes
Public	Glenn Review April 1994	Forward Estimates Strategy Papers (FESPs)	Finance	Development of policy papers

Release	Uhr Review June 1995	Role of Finance in Promoting and Using Evaluation	Finance	IDC policy processes
	Weller Review June 1995	Commonwealth-State Reform Processes (COAG)	DPMC	IDC policy processes
Confidential	Argy Review May 1993	Policy Advice on Government Response to the Martin Inquiry	Treasury	Development of policy papers

Having settled on process as the focus for the evaluation of policy advising programs the immediate concern was determining *what* to evaluate. Often departmental structures and program responsibilities overlapped rather than coincided. In these circumstances, the processes of policy advising could be interpreted in a number of ways; it could refer to a review of the performance of nominated organisational units, the management of resources with respect to discrete policy tasks, or a combination of both in terms of structures that are specially created for managing policy development processes, such as IDCs. IDCs are, of course, the traditional bureaucratic coordination mechanism, since, by definition, they deal with policy problems that affect a number of departments, while at the same time providing “arenas for the meeting of nominal equals” (Painter & Carey 1979, pp1-3). If PMRs were to remain faithful to the intention of *program* evaluation—and hence assess discrete policy advising tasks—then the most amenable way to do this was to concentrate on IDC processes.

However, before proceeding any further, we should first acknowledge some of the possible limitations in directing (the majority of) PMRs to policy advisory processes framed around IDCs. For instance, one could claim that the PMR format is not likely to tell central agencies much about advising processes that are not already known to them. The problems of IDCs as instruments for policy development and coordination have been comprehensively revealed in a number of academic studies (eg Painter & Carey 1979; Self 1980). Are these identified problems intractable, and hence beyond the policy management capabilities of central agencies, or can they be tempered? Alternatively, is the PMR format limited in its explanatory power *because* it confines itself to IDCs which serve particular purposes? In this case, can it be claimed that the IDC process is fairly representative of the advisory and coordination functions of central agencies? On these last two points, we can note from earlier consideration that a large proportion of central agency work is either reactive or broker-like, and often covers both cross-portfolio and cross-program issues. Under these circumstances, it is reasonable to assume that IDC processes are an established norm of central agency work.

The approach taken here to analysing the results of PMRs is to pose a simple question: What do PMRs tell central agencies about the processes of policy advising? Each of the reviews accepted that their primary focus was on the process of policy advising. At the same time, it was also recognised that this focus would only be of significance to central agencies if it concentrated on the *effectiveness* of those processes, and, further, that the effectiveness of advisory processes could not be assessed in isolation from the policy advice product, whether it was a policy options paper, a specific strategy paper or broad intergovernmental agreements. As an instance, take the Glenn Review of the 1992-93 Carers Package, an IDC that considered options for the expansion of services and financial assistance provided to carers (Glenn 1993). The objective of the review, as stated in the Review, was to

> assess the effectiveness of the Department's [DPMC's] policy development processes in an important area of activity and ascertain whether improvements need to be made to its role in developing policy advice for the Government (Glenn 1993, para2.3, p8).

More specifically, the Review was structured around the "process followed and the outputs from the process including the linkage to decision making" (Glenn 1993, para 1.5, p2). Somewhat more methodically, the Weller Review adopts a similar, although cautiously circumscribed approach (Weller 1995; an abridged version appears in Weller 1996). In examining the management of Commonwealth-State reform processes that serviced the Council of Australian Governments (COAG) initiative, the terms of reference provided for the evaluation of "organisational arrangements and processes adopted . . . with a view to identifying the key elements of good policy development" (Weller

1995, p1). It was conceded by the reviewer that it would be "easier to assess what processes work than who made essential contributions", and so a formal analysis of policy stages was undertaken so as to "determine at which point difficulties are met and what has been achieved" (Weller 1995, p2).

From this brief introduction to the objectives of just two PMRs, we can see that the reviews, which seek not only to describe and assess the way central agencies have undertaken policy advisory tasks, but also offer prescriptions on how they these might be improved, are characterised by a number of common elements. The contention here is that there are three primary areas of policy process assessment that PMRs seek to address—the management of the policy process, the appropriate role of central agencies in policy development and coordination, and, in a tentative fashion, the assessment of the influence of advice on decision makers. The last of these can be seen as a "spillover" attempt to evaluate policy advice output. Each of these will now be examined in detail.

The Management Of The Policy Process

The management of policy process has emerged as the leading area of concern for PMRs. This can take one of two forms. The first is basically an analysis of the limitations of the IDC process for generating good policy advice, while the second converse task is to identify those practices that are crucial for successful policy development. Without exception, those reviews concerned with IDC processes confirm the criticisms sourced from academic studies. They indicate that the defining characteristics of IDCs—a tendency for departmental representatives to act first as delegates defending departmental territory, in conjunction with the practice of searching for a consensus outcome so that representatives can use IDCs to delay decisions detrimental to their department—border on the intractable (Painter & Carey 1979, pp62-63). One consequence of these attributes is that assessing the effectiveness of IDC policy processes is a formidable task. For example, the most obvious measure of the quality of IDC policy processes might be the extent to which policy recommendations coincided with what the government actually wanted. However, the criterion of ministerial satisfaction is itself imprecise, since more often than not IDCs are established precisely because in certain policy areas ministers *do not know what they want*—policy development is displaced to the bureaucracy (Painter & Carey 1979, pp4, 17-21, 33, 36-38, 42-45, 66, 98-103). In this way any gauge of IDC policy effectiveness could be verging on both the self-referential and the self-defeating. Nonetheless, throwing a light on IDC processes does, however, reveal much about the quality of advisory processes.

A number of these issues gained prominence in PMRs, particularly the need to reduce the influence of departmentalism and ensure a balanced IDC composition. The Glenn Review of the 1992-93 Carers Package provides an instructive example. In this case an IDC, which was chaired by DPMC, was commissioned by Cabinet to develop options for expanding the services and financial assistance provided to carers of ill, frail and disabled dependents “against the expectation of new funds being available”, although this expectation was later to dissolve (Glenn 1993, para 7.1, p23). The review was most complimentary to the DPMC in this process, describing its role as

> decisive. It provided the leadership required to get the result, adjudicated on disputes, ensured rigour was applied and effectively co-ordinated and managed the process without being possessive or intrusive. It clearly demonstrated the honest broker role that only a central agency can play in bringing together parties with competing interests, reaching common ground and producing a report for Government (Glenn 1993, para 1.10, p3, para 6.30, p22).

However, the review also noted two major deficiencies in the IDC process that hampered its progress. The first related to the rampant pursuit of sectional interests within the IDC that made concrete agreement on the options to be submitted to ministers an unrealisable objective. Discussion was impeded by the “entrenched interests” of departments, and compromises had to be struck. In the end a total of twenty-three options, placed into three policy packages, were developed (Glenn 1993, para 1.16, p5; para 7.6, p24). As a consequence, the review was of the opinion that the

> outcome which flowed from the pursuit of sectional interests by IDC members was effectively the lowest common denominator. It meant that too many options were submitted to Cabinet without the priorities being sharply enough defined for Ministers (Glenn 1993, para 7.14, p25).

Here then is a classic case of an IDC searching for agreement that intentionally avoided conflict, despite the best coordinating efforts of DPMC.

The effects of departmentalism were compounded by two further weaknesses of the IDC process. Representation on the IDC was dominated by social welfare agencies. The absence of an advocate of economic or financial positions, other than the Department of Finance, that could have provided some “balance” to the negotiations contributed to the protracted decision making process (although this reflects as badly on Finance’s responsibilities for advising on expenditure policy) (Glenn 1993, para 6.14, p18). The IDC process was also impeded by the fact that the Government had not provided any “specific guidance” on the direction that policy should take. As a consequence, progress of the IDC was further frustrated by a need to determine its own terms of reference and policy objectives (Glenn 1993, para 6.8, p16; paras 7.2-7.3, p23). Again both of these points are well-established marks of a poorly managed IDC process.

In the Carers Package IDC the first Glenn Review found many of the standard problems confronting IDC policy processes which were, apparently, resistant to the policy management capacities of DPMC. At the same time, it could also be said that their ill-effects may have been more pronounced if not for the “honest broker role” played by DPMC. Both interpretations are equally plausible, although the negative case appears stronger on the evidence provided in the PMR. Nevertheless, the constructive aspect of the review is that it provides some guidance on how better to structure IDCs. However, what of other reviews, and policy management practices deemed crucial for good process?

The Weller Review of DPMC’s management of Commonwealth-State reform processes (the COAG forum) found many of the same IDC practices done well.[7] In this instance, DPMC chaired the Commonwealth-State Officials Steering Committee (OSC) which serviced the COAG Heads of Government meetings. This servicing role was pivotal, since the OSC recommended policy options,

[7] Designating COAG processes as IDCs is not unduly stretching the definition of an IDC, as the Weller Review notes (Weller 1995, p5). COAG structures, admittedly, represent a special kind of IDC, but nonetheless hold the essential characteristics, *viz*. heads of government (and their officials) entering the process as equals, with the consequence that “the balance of power in negotiations is more even” than that presiding at the Special Premiers’ Conferences they replace.

secured their place on the COAG agenda and advised on priorities. Working across two levels of government complicated the IDC process, but DPMC's response, according to the review, was tactically to manage the process with professionalism and technical competence (Weller 1995, p17). Many of the obstacles of the IDC process were avoided by astute policy process management—carefully introducing policy issues in broad terms that did not threaten particular interests, securing endorsement first and then concentrating on policy details; ensuring IDC progress by assigning clear responsibility for discrete tasks to specified individuals; ensuring continuous ministerial involvement; and deciding early what consultation was required, all the while bearing in mind that "consultation [was] not negotiation; it is seeking comment, not necessarily endorsement" (Weller 1995, pp9-10).

This last practice of consultation attracted universal approval. Ensuring that most, if not all, relevant interests are considered when developing policy initiatives was, not surprisingly, ranked very highly by all four PMRs as an imperative for both demonstrating the propriety of the policy process, as well as increasing the likelihood of achieving policy outcomes desired by central agencies. Effective management of the policy process calls for consultation of interests both internal *and* external to government—the absence of genuine consultation can seriously undermine policy development. For instance, both Glenn Reviews noted how insufficient consultation could narrow the policy choices available to advisers. The Carers Package IDC was hampered by a reluctance on the part of DPMC to approach State Governments for their perspective on providing assistance to carers (Glenn 1993, paras 1.18, 1.20, p5). Similarly, the Glenn Review of FESPs, directed at assessing the Department of Finance's role in preparing forward estimates strategy papers (FESPs) for Cabinet, was also critical of that department's lack of consultation, this time with line departments and their ministers. Designed to inform budgetary decision making, FESPs outlined options for containing outlays growth in the big spending line departments. In this, they relied heavily on cooperation from other departments for program expenditure data. The review found, however, that consultation was negligible, confined as it was to checking factual material. This meant that savings options were prepared in isolation from, and then sprung on, line departments, not only slowing paper preparation but also causing unnecessary resentment (Glenn 1994, para 1.11, p4). This was judged a "mistake" (Glenn 1994, para 7.6, p21).

On the other hand, while some consultation is preferred to none, it must always be managed with a firm hand. The Uhr Review demonstrates this effectively (Uhr 1995). This PMR examined the involvement of the Department of Finance on two pre-budget IDCs directed at reviewing policy in the disparate areas of aged care services for Aboriginal and Torres Strait Islander (ATSI) groups, and cost recovery practices in Commonwealth management of Australian fisheries. Consultation was important to the conduct of both IDCs—and successfully applied in the case of the ATSI IDC—but for the cost recovery IDC there was some evidence of attempted interest group capture. In this case, when one interest group was given standing on the IDC, Finance faced the threat of having the IDC locked into an unsatisfactory policy position prior to final budget Cabinet deliberations (the fishing industry body claiming that Finance had sought to amend agreed IDC policy decisions by having only the Finance position recognised in the final report) (Uhr 1995, paras 6.8-6.10, p25; paras 6.14-6.15, p26). The lesson was clear: consultation is a mechanism that must carefully be managed by central agencies within the IDC context so as to inform, but not dominate or capture proceedings (cf Weller 1995, p10).

Determining The Appropriate Role Of Central Agencies In Policy Development And Coordination

In addition to analysing policy advising processes in order to identify both the pitfalls and the paths of least resistance, PMRs have also addressed the quite fundamental question of what should be the appropriate role of different central agencies, most obviously in managing the IDC policy process. This is an issue that goes to the core of the policy management function. In the course of managing policy advisory processes, are the separate dimensions of policy management, such as resource coordination and policy development and coordination, accurate descriptions of the work of individual central agencies? Alternatively, do the categories hold any prescriptive value in marking out the limits of central agency capabilities? It is argued that the trial series of PMRs, through the assessment of the effectiveness of IDC advising processes, confirm important aspects of the policy management roles of DPMC and the Department of Finance.

In sum, the reviews suggest that there are principally two levels at which central agencies exercise policy management. The first operational level describes *how* policy management functions should be exercised under different IDC conditions. The second framework level marks out the core responsibilities for individual agencies when managing the policy process. It describes the limits of their policy competency. Each of these will now be examined in more detail.

Earlier, mention was made of the role played by DPMC during the 1992-93 Carers Package IDC. The review saw this role as requiring leadership, adjudication skills and a capacity to coordinate and manage the process "without being possessive or intrusive" (Glenn 1993, para 1.10, p3). In effect, a brokerage position was taken in what turned out to be a task of conflict management. It is the nature of this brokerage role that has caught the attention of other PMRs. For instance, in examining the contribution made by the Department of Finance to two very different IDCs, the Uhr Review emphasised how the origin, purpose and composition of the IDC was likely to determine the role played by participating central agencies. While noting that both IDCs had been established by Cabinet ministers, each reflected different origins (the ATSI IDC was a response to sharp public criticism, while the cost recovery IDC was a policy review in response to a government commissioned report) and distinct compositions (the ATSI IDC was large, replete with representatives from all relevant agencies, but involving Finance only at the behest of DPMC, while the cost recovery IDC was small and contained within the relevant portfolio) (Uhr 1995, paras 4.3-4.4, p12). The review adjudged that it was these differences that influenced the way Finance approached its policy management tasks. For the former IDC, Finance adopted a "utility player" role where intervention was minimised, while in the latter case, where Finance was the only agency from outside the affected portfolio, a more "active" role was required to reach an acceptable result (Uhr 1995, para 4.5, p13; para 5.8, p17).

A similar conclusion was reached by the Weller Review of COAG processes where, as we have seen, DPMC's task revolved around the defining of policy issues and the development of workable policy agendas. In this, DPMC acted as either a "broker" between departments, as a "facilitator", or as the "direct source of policy ideas"; that is, it seemingly exercised both reactive and proactive policy advisory roles (Weller 1995, p12-13). However, the review argued that the bifurcated nature of the COAG process demanded that DPMC "participate actively" in two complementary ways (Weller 1995, p15). The first entailed maintaining progress and consistency *within* the Commonwealth Government by breaking policy deadlocks between portfolios—it was responsible for "delivering" Commonwealth line departments at the negotiating table. The second task was to ensure the vitality of Commonwealth-State relations by facilitating negotiations conducted through the OSC. Again, IDC form influenced central agency responses.

Observations of this variety sit comfortably with the distinction made earlier in this chapter between general and specific policy advising. The reviews suggest that in managing the advisory process, central agencies are likely to be faced with a choice between assuming either an active or reactive posture, although it appears difficult at times to distinguish clearly between the two. In deciding when to stress and relax the respective roles, central agencies should, at the very least, consider the type of IDC conditions facing them. But these considerations will also be constrained by the other level of policy management, core responsibilities. This is most clearly demonstrated for the Department of Finance.

The Uhr Review takes a special interest in the core responsibilities of the central budget agency. There are, according to that review, clear limits to what contribution the Department of Finance can and *should* make to the IDC policy advisory process. The argument runs as follows. As we learned in Chapter Six, Finance's resource coordination function rests heavily on a capacity to evaluate programs against "value for money" criteria. However, as the Uhr Review does, a plausible distinction can be made between "assessment" (which is a determination characterised by impartiality) and "evaluation" (which, in addition to impartiality, calls for "subject matter expertise") (Uhr 1995, paras 3.11-3.13, pp9-10). But in claiming proficiency in "evaluation", Finance can only point to one of two capabilities: either that it possesses subject matter expertise across government policy, or "that it has specific and distinctive expertise in some vital aspect of government policy that is unshared by other agencies". That distinctive field is advising on expenditure policy, not the substantive policy field of a particular agency (Uhr 1995, para 1.6, pp2-3; paras 3.15-3.17, pp10-11; paras 7.2-7.5, pp29-30). Hence, the core responsibility for Finance involves assessment (rather than evaluation) on issues of expenditure policy and financial management practices, leaving general policy development and coordination to other central agencies (presumably DPMC).

One implication of this prescription might be that resource coordination and policy development and coordination are two distinct dimensions of policy management that will rarely, if at all, intersect. As general guidance it has much to commend it, since sustained claims to subject matter expertise on the part of a budget agency could expedite the conditions for “policy overstretch” (see Chapter Three). At the same time, however, this interpretation also underplays the fact that quite often expenditure issues and substantive policy issues are one and the same. Determining value for money will mean making judgments about the relative merits of different substantive policies (on which refer Chapter Six). The task facing a budget agency is to *judge* where and when this “assessment” role applies, and where and when “evaluation” does not. If anything, determining core responsibilities will, like the exercise of policy management functions, be somewhat contingency based. In the end this is heavily dependent on professional expertise, although PMRs may be one instrument for better informing the application of this expertise.

Spillover Effects? Assessing The Influence Of Advice On Decision Makers

It does come as something of a surprise to learn that issues relating to the influence of policy advice on decision makers were canvassed in some of the PMRs, most notably the two reviews conducted by Glenn which expressly included the aim of assessing the effectiveness of policy advice products for ministers and their advisers (Glenn 1993, para 1.5, p2; 1994, para 1.3, p2, para 2.3, p6). As we have discussed, one of the chief problems facing this form of assessment is that it enters what many central agency officials consider to be a restricted zone, where reviews would not only be judging the quality of the policy advice product, but, in some cases, commenting on the content of policy, which properly was the province of ministers. This, apparently, was the case with the Argy Review undertaken for the Treasury and, as senior Treasury officials explained, was the prime reason for its non-disclosure. Similarly, the Glenn Review of the 1992-93 Carers Package had a delayed public release, in its entirety at least, until mid 1995 when its availability coincided with the release of the Working Group’s report. Its attempt to scrutinise the effectiveness of policy advice outputs was in all likelihood a contributing factor.

Nonetheless, the Glenn Review does provide an instructive example of a foray (that was not specifically sanctioned by DPMC) into the effectiveness of policy advice outputs. In considering DPMC advice that was provided *after* the IDC memorandum had been submitted to Cabinet, the review notes that the DPMC briefings on the Carers Package recommended a package different to that of the IDC memorandum (Glenn 1993, paras 8.2-8.5, pp29-30). This was not surprising given the outcome of the IDC. However, although the review noted that the DPMC briefing "hit the right target areas", DPMC had also persisted with developing policy options "which seemed to have little support among Ministers" and failed to give "more serious consideration to savings options when ministers were seeking them" (although this narrowness was also attributed to a working assumption provided by the Government of increasing outlays) (Glenn 1993, para 1.22, p6; para 7.12, p25; paras 8.6-8.7, p30). The review concluded that the diffuse nature of IDC and DPMC policy recommendations both unnecessarily and unsatisfactorily made extra work for ministers. Without explicitly saying so, the review suggests, in two separate instances, that DPMC had failed to give the advice that ministers actually sought:

> That Ministers themselves had to find the way to effect the savings does not reflect well on the advisers generally (Glenn 1993, para 8.8, p31).
>
> The final decision taken by the Government was, therefore, broadly in line with the advice given, but the process was drawn out. To get the key proposals and expenditure worked out took many weeks and required a good deal of attention from Ministers (Glenn 1993, para 8.11, p31).

For both DPMC and Treasury this might constitute the realisation of mutually held fears about publicly assessing the effectiveness of policy advice in terms of its effects on decision makers (that is, ministers). This, however, is clearly not the case. As the extracts above demonstrate, the final advice products—that is, options papers from both the IDC and DPMC—were judged with explicit reference to the quality of the processes followed during the entire IDC process. Even though DPMC's performance in coordinating other departments during the IDC was rated as "excellent" (Glenn 1993, para 6.27, p21), the deficiencies of the overall process produced a policy advice product that did not, in the first instance, satisfy the demands of ministers.

Analysis of this type raises a number questions about the suitability of PMRs for assessing both the process and output of policy advising.[8] Did the recommendations simply not reflect what the Government wanted, regardless of analytical rigour? Alternatively, were the recommendations more a reflection of those policy positions preferred by DPMC? At what point should policy making end with ministers and start with departments? It is somewhat ironic that the review's concerns about ministerial involvement being forced by IDC procedural inadequacies are precisely what others have diagnosed as the key remedy for the maladies of IDC processes—in order for the political executive to maintain control of policy development, the collective power of ministers should be accentuated through *increased* ministerial involvement in IDC processes (Painter & Carey 1979, pp95-116).

The Treasury experience with PMRs is a marked contrast to (what one can only presume were) the relatively innocuous reviews conducted on behalf of DPMC. From the interview evidence, it appears that the Argy Review (Argy 1993), which remains confidential, was the only PMR that sought to evaluate the quality of policy advice *as well as* comment on the appropriateness of the advice, rather than confine itself to process. For this reason it was buried, since it was judged by the Treasury to be too sensitive politically. The Argy Review was undertaken *in camera* and examined material prepared for the Government's policy statement in response to the Martin Inquiry into Banking. Within the Treasury the review was castigated as either inadequate—"it reads more like a discussion paper than management assistance" (Interview: 01)—or "useless", since it explicitly commented on the content of government policy, which was outside its scope, rather than the policy advising function (Interview: 27). Although it was readily conceded that the reviewer did not

---

[8] Similar effectiveness issues were explored in Glenn's second review of Finance's preparation of FESPs for the 1993-94 Budget (Glenn 1994). The main problem here was the level of uncertainty surrounding the objectives of the strategy papers. Users, comprising both ministers and advisers, were "confused" as to the primary objective—were FESPs merely to outline what was driving outlays growth and list policy options, or, instead, provide a detailed analytical brief? Excessive attention to achieving consistent presentation displaced the quality of content (which erred on the side of detailed analysis) with the common view amongst users being that they were swamped with information, and that this was "making it difficult for Ministers" (Glenn 1994, para 5.6, p13; paras 6.10-6.12, pp18-19). The quality of the advice product was again assessed in terms of process; in this case, inadequate attention to the needs of the user for concise and timely advice.

receive a clearly defined brief (Interview: 21, 27), the episode has confirmed Treasury's suspicions about the subjectivity of the PMR format.

Evaluation Fatigue? PMRs And Policy Management

The analysis above confirms that the PMR format motions a return to managing the policy process at the centre of government, with reference to the advice product but not so as to frame assessment around policy outcomes. As the final phase of the evaluation strategy, it also represents something more. We have noted how the extension of effectiveness evaluation to policy advisory programs was a response to mounting external pressure for central agencies essentially to practice what they preached. In this way, the PMR format was merely taking program evaluation to its logical conclusion—if, under the evaluation strategy, line portfolios could be held publicly accountable for the effectiveness of their service delivery and regulatory programs, then surely the same standards must also apply to the policy advice generated by the central agencies? At the same time, however, the entire PMR enterprise was also displaying signs of increasing fatigue on the part of central agencies in adhering to the tenets of program evaluation.

As discussed earlier, the PMR format represented a compromise position that settled on evaluating the process rather than the results of policy work. In policy advising programs the evaluation strategy had found the limits of evaluating program outcomes, and the boundaries proved not only to be methodological but also conceptual. It was clear from the development of the PMR format that one of the major underlying tensions of program evaluation—which had been acknowledged some twenty years earlier in the report of the RCAGA—had reappeared. This tension hinged on the determination of public service responsibility for government policy. By clearly distinguishing between ministerial responsibility for policy outcomes and public service responsibility for policy process, the Working Group had, through the PMR format, marked a conspicuous departure from the evaluation strategy's preoccupation with evaluating both the effectiveness *and* appropriateness of government activities. The return to process was a retreat from any increased public service responsibility for making policy back to the conventional divide between policy and administration.

In this, the development of the PMR format—and, in particular, the strong opposition from the Treasury—was replete with the caution outlined in the *Report* of the RCAGA, that program evaluation could *not* serve public accountability purposes precisely because it dealt with issues of policy effectiveness. For the RCAGA, program evaluation (which, in its view, would "rarely be possible" (RCAGA 1976, para 3.3.13, p41) for policy advising programs) should be designed solely as a central coordination tool internal to government. As we learned in Chapter Four, despite the RCAGA's attempts to articulate a new public service responsibility for policy effectiveness, the policy-administration dichotomy (or its contemporary equivalent, the efficiency-effectiveness divide) formed a ready made barrier to any corresponding increase in public accountability. In a limited way—by suggesting public service responsibilities for program effectiveness and appropriateness—the evaluation strategy was crossing this divide. The PMR format then provides some evidence that the strategy was becoming fatigued. Not only did it depart from this outcomes focus by returning to process, but of the two original purposes credited with driving policy advice evaluation (internal policy management and external public accountability) the first has seemingly eclipsed the second. This can be directly attributed to the tension just outlined. For accountability purposes the PMR format attempts to separate process from output and, in so doing, excise policy or "politics" from the policy advisory process. The problem, as one PMR reviewer has observed, is that political considerations can not be quarantined from either the advisory process *or* the product when conducting evaluation. The issue is whether these considerations can or will be included:

> Policy advice is essentially a political commodity, and is most appropriately assessed in political terms . . . To pretend otherwise, and to experiment with so-called independent and external evaluations of policy advising, is to pretend that bureaucratic advisers are somehow beyond politics . . . (Uhr 1996, p105; cf Waller 1996, p72).

The onset of fatigue in the application of evaluation to policy advising programs has finally acknowledged the *external* political constraints facing program evaluation at every level.

The analysis of PMRs undertaken earlier suggests they have unfolded as guidelines for increasing the influence of central agencies over the processes of policy advising. Policy management, as envisaged by PMRs, becomes a task of reducing conflict by finding the paths of least resistance through the messy IDC process. The most effective way to do this is for central agencies to tighten their management control of the advisory process to ensure that they are not dominated by other government (or, for that matter, non-government) interests. PMRs as an instrument for enhancing central agency management of the policy process hold considerable currency at the most senior levels of all three participating central agencies. One deputy secretary within DPMC has, in a number of quarters, strongly pushed this line, arguing that one of the most significant yet neglected aspects of the role of central agencies is "how to handle strategically the policy process" and manage the potential conflict both within and between departments to new policy initiatives (Edwards 1992, p448; Edwards & Stuart 1994, pp4-7, 18-19; Edwards 1996, pp73-74). In this way the

> focus on the policy content is only half the story; it is also necessary to assess advice given on how to carry proposals through bureaucratic as well as political processes in order to get desired outcomes (Edwards 1992, p448).

There is then, a receptivity within the central agencies to PMRs as "guidelines" for managing policy development processes, particularly of the IDC variety. To this extent the results of PMRs have, according to a number of senior officers, already proved to be of use in DPMC for the recording and retention of corporate memory, as well as a ready made resource for training in policy advising by providing "good rules of practice" (Interview: 08, 12, 15; Edwards 1996, p79; Gregory 1996, pp143-145).

At the same time adherence to the external purpose of public accountability has waned, and the views of senior public service management, divided as they are, reflect this. In the one corner are the senior executives of both DPMC and Finance who readily subscribe to the "transparency" school (Record of Conference Discussion in Uhr & Mackay 1996, p130; Interview: 08, 12). In this case the PMR must be treated like any other program evaluation and "should be *put in a form* which allows it to be published". The italicised words are of particular interest, since both secretaries reserve for individual agencies the right to produce two "versions" of a PMR, a public version, and a confidential version that includes the classified material that may have been used to reach the conclusions (Record of Conference Discussion in Uhr & Mackay 1996, p136).

In the other corner are the ever vigilant sceptics in the Treasury Department. In this locale, PMRs are viewed as eminently more appropriate for internal departmental use "in order to improve the Department's own performance, rather than to meet external accountability concerns" (PSC 1995, p28). According to senior Treasury officers this position is generated by two developments (Interview: 21, 27). The first revolves around Treasury's experience with the Argy Review, which has precipitated considerable disenchantment with the review process (colourfully described as a case of "once bitten twice shy"). The second reflects an uncertainty about fundamental public accountability relationships—to whom are public service advisers primarily accountable, the executive or the Parliament? Some Treasury officers found the accountability side of PMRs to be a "huge problem" since, to them, it is not clear who the Treasury works for. If the primary client is the executive, then evaluations must be held *in camera* since there are statutory limitations on the public release of policy advice. Echoing the reviewer concerns outlined earlier, in this situation the criteria for evaluation probably needs to include what one officer termed *political effectiveness* (Interview: 27).

The PMR enterprise was principally an attempt by central agencies to preserve the credibility of the evaluation strategy both within the public service and the wider community. The aim was essentially consistency in application. Convenient as it was for the central agencies, the extension of their influence over the processes of policy advising is perhaps best seen as an unintended by-product of the PMR development process and the old constraints that were heightened by evaluation fatigue. As a compromise position, the PMR format represents a trade off between public accountability and internal utility. It also represents something of a vindication of the reservations contained in the Treasury Department's position.

A final question then, is whether PMRs could ever adequately perform an accountability function in terms of policy advising work. Parliament seems barely to be cognisant of their existence, let alone employing them as evidence of stewardship in the provision of advice and carriage of policy development. There are also signs that PMRs may not always provide publicly credible performance information. The sanitised "public" versions of PMRs will inevitably exclude much of the classified evidence of either good or bad advisory processes. Further, there is every likelihood that the advising programs or the discrete policy tasks to be evaluated will be carefully selected so as not to reflect too badly on the department's performance.[9] There is certainly an unarticulated concern on the part of Parliament with the adequacy of policy management processes generally (see Uhr & Di Francesco 1994, pp249-251), and given that efforts to evaluate policy advice were initiated by parliamentary interest, then the onus may lay with that institution to ensure that an accountability role for PMRs retains viability.

CONCLUSION

The evaluation of policy advice is a component of the evaluation strategy that remains in a state of flux, but that, in the main, has emerged as an instrument by which central agencies can extend their influence over the processes of policy advising. This chapter has made a case for two major arguments that bear this assessment out. The first contends that, because of its focus on process, the evaluation of policy advice represents a significant departure from the dominant outcomes framework of the evaluation strategy. It has shown that the PMR format was the product of bureaucratic expedience—after an early reluctance to extend the evaluation strategy to policy advising programs, the evaluation of the effectiveness of advisory processes was shaped by a combination of methodological constraints and strong bureaucratic opposition.[10] The second contention is that the purpose of PMRs has shifted from what was primarily external pressure to develop a public accountability mechanism for policy programs, toward an internally oriented instrument for

[9] In this respect several reviews have noted the limitations of the PMR methodology, especially for public accountability purposes. In the main these relate to restricted access to select documentation, obligations of non disclosure, and the absence of incentives for officers in line departments to criticise the work of central agencies (Uhr 1995, pp4-6; Weller 1995, p15).

[10] It also accords with the conventional "wisdom" on managing the politics of policy advice that places great store in the advisory process as the crucial pressure point for rulers (Ministers). However, not only does the onus rest with the ruler to "calibrate" the advice they receive, but the evaluation of policy advice must be a part of the process of advising in order to ensure good advice (Meltsner 1990, pp116-126). One question which remains open, of course, is what role the *political* executive (not to mention wider community groups) have actually played in pushing the evaluation of policy advice. This chapter suggests it has been a minimal one—the PMR format was designed by and for the bureaucracy.

augmenting the policy management functions of central agencies. An analysis of the PMRs that are available suggests that this assessment tool holds considerable potential for increasing the leverage that central agencies have in the IDC policy process, as well as providing some guidance on the limits of their core policy management responsibilities. In this way, by returning to policy process, the evaluation of policy advice represents the most express application of evaluation to policy management purposes.

Chapter Eight

Ploughing The Policy Sands? Evaluation And Policy Management

This thesis set out to chart the development of the first policy of systematic program evaluation within Australian government administration, as well as assess its effectiveness against what emerged as its primary objective, informing the policy management functions of central agencies. The evaluation strategy was, by international standards, a highly credible attempt to institutionalise program evaluation within government. Born into an environment of acute fiscal stress, the distinguishing mark of the strategy was its attempt to explicitly promote substantive *policy* considerations, rather than simply financial concerns, in priority setting by government decision makers. To this end, as we have seen, the strategy was admirably comprehensive in its scope. The real merit of program evaluation was that it aimed to raise the general level of awareness within the administration that the overall effectiveness and continuing relevance of government activities should matter in the calculus of government decision processes. Determining the extent to which it did matter was an important part of the brief of this study.

The task then of this concluding chapter is twofold. First, it draws together the findings of the two case studies contained in Part Three and examines the implications of the evaluation strategy for central agency policy management. The contention is that the influence of program evaluation has been confined to marginal aspects of both the budgetary process and policy development and coordination, and that the level of enduring change in either has fallen considerably short of that anticipated by the strategy (although the final impact on the latter remains to be seen). Second, it builds on the evidence adduced in earlier parts of the study, and presents an analysis of the limitations of the evaluation strategy as policy design. Principally, it points to the uncertainty of political choice as the crucial factor conditioning the institutionalisation of any policy analytical technique.

PROGRAM EVALUATION AND POLICY MANAGEMENT

As a point of departure let us first review the two main arguments of the thesis. The first is that program evaluation policy has been directed at questions of coordination in central government, principally those relating to the central budgetary process. A detailed historical account of the development of evaluation policy since the 1976 *Report* of the RCAGA described a highly competitive policy arena in which a number of institutional interests sought to define the bounds of the new criterion of policy review—effectiveness. Program evaluation was variously seen as serving parliamentary concerns for increased public accountability, central agency requirements for policy coordination, and portfolio and community based interests for more responsive program management. Although traces of all three design characteristics could be found within the evaluation strategy, the dominant function was evaluation as an instrument for central agency policy and resource coordination. This objective was cemented by the onset of both acute fiscal stress and extensive economic and public sector restructuring during the latter half of the 1980s and the early 1990s.

However, both economic *and* public sector restructuring were, to some extent, tempered by the Labor Government’s policy concerns for social justice. In this way, the evaluation strategy followed two agendas. At one level the amalgam of objectives sought to accommodate the competing interests in program evaluation by according them equal standing. The evaluation strategy was an integral component of financial management reform across the Australian bureaucracy and this agenda served to market the strategy as “evaluation for all seasons”. At another level the central agency responsible for the evaluation strategy, the Department of Finance, clearly pursued program evaluation as a tool for enforcing the government’s agenda of decremental budgeting, although a tool that was clearly angled towards the effectiveness and relevance of policy rather than economy. Not only would this address the Labor Government’s pressing need for targeted fiscal restraint, but it also secured the position of the department during a period of volatile public sector restructuring.

In light of these developments, the second argument of the thesis is that the evaluation strategy can best be examined as an attempt by central agencies to enhance what are here referred to as their policy management functions. This function is framed around the management of the coordination problem in central government. Program evaluation is one response to the transformation of the coordination problem from an initial concern with the process of government to a sharpened focus on substantive policy concerns (such as the effectiveness and continuing relevance of programs). The thesis confirms that program evaluation was directed at augmenting both the resource coordination and policy development and coordination dimensions of central agency policy management. The influence of evaluation, however, was found to be marginal, particularly in terms of central agency management of the budgetary process. In addition, the extension of evaluation to policy advisory programs marked a return to the procedural concerns of coordination and a departure from the policy effectiveness orientation of the evaluation strategy.

On the basis of these general findings, and in terms of the narrow focus of this study on central agency policy management, to what extent can the evaluation strategy be adjudged a failed excursion into the realms of formal program evaluation policy? In order to address this fundamental issue we must return to the two questions posed in Chapter Three, and examine the implications of the study for the policy management capacities of central agencies in terms of both budgetary decision making and aspects of policy coordination.

Question One. To what extent has program evaluation led to enduring change in the rules of policy management?

This first question was essentially directed at the nature of the tasks that central agencies faced and the extent to which program evaluation may or may not have altered central budgetary and policy coordination processes. It can be recalled from Chapter Three that the comparative literature which focused on the expenditure budget process indicated that efforts to strengthen the capacity of budget agencies to exert influence in resource coordination were commonly hampered by the weakness of the link between a program evaluation policy and budget review procedures. More significantly, the influence of formal budget review procedures and techniques were often shaped by the informal networks of bureaucratic relationships—formal policy analytical techniques failed to recast the rules of the policy management 'game' and instead were absorbed into the existing framework of rules and relationships.

Based on the findings relating to program evaluation and resource coordination presented in Chapter Six, it is safe to conclude that the evaluation strategy has suffered a similar fate. This chapter focused on how aspects of the budgetary process are now conducted following the introduction of the evaluation strategy in late 1988. It presented two sets of evidence. The first examined the impact of program evaluation on the work of budget agency officials. As an instrument for advising both on the merits of policy submissions and the location of possible savings measures, the evaluation strategy has proved difficult to manage. The evaluation planning process was, by design, heavily weighted against the central budget agency, the available performance information was either not suited to cabinet level priority setting or was deployed by line departments to act as a shield for programs, and the evaluation planning cycle did not coincide with the overriding budget cycle. In summary, either the rules of the resource coordination game have witnessed little substantive change, or the players did not recognise the change that had occurred, for there remained within the expenditure review branches of the Department of Finance significant pockets of scepticism and resistance towards program evaluation.

The second body of evidence, which was based on utilisation data collected by the Department of Finance, indicated that while evaluation of some type *was* being referred to in the preparation of both Finance and portfolio budgetary advice, it was not on the systematic basis intended by the evaluation strategy. Program evaluation was referred to on an *ad hoc* basis, when it was available. Further, there was little evidence to suggest that the evaluation that was relied on was *more* oriented towards how effective programs had been in achieving their objectives. At the same time, the political executive, the ultimate user of program evaluation, displayed signs of dissatisfaction. Evaluation was part of an exercise in expenditure restraint, but these objectives, particularly towards the end of the fourth and fifth Labor Governments, were not served as anticipated. As a result, sharper instruments in the form of purpose built program reviews were introduced.

The budgetary use of program evaluation highlights how important the *informal* aspect of review is in the work of the central budget agency. Program evaluation was an attempt to formalise aspects of the analytical review work of the expenditure review branches of the Department of Finance. A crucial problem appears to be that these formal aspects are heavily dependent on the informal processes, which comprise the judgment, intuition and experience of supply officers in identifying emerging policy priorities, suspect programs that may require investigation, as well as possible targets for savings options. And, in turn, these abilities are deeply rooted within the networks of personal relations that budget officials maintain throughout the bureaucracy, as well as the deftness with which they conduct budget negotiations (cf Heclo & Wildavsky 1974). The way in which the evaluation planning game relied on the bureaucratic cunning of both expenditure review officers and portfolio officers illustrates this. On reflection, the evaluation strategy could be interpreted, at the budgetary level at least, as having laid its foundations in sand—it was based in a weak capacity to formalise processes that are not, due to the nature of the work, amenable to formalisation.

In the case of policy development and coordination, which was examined in Chapter Seven, the evidence is less convincing and the conclusions somewhat tentative, if only because the relevant evaluation instrument—PMRs—remain at the experimental stage in terms of both conduct and application. The evaluation of policy advice proved to be a revealing test of the endurance of the evaluation strategy, since the PMR format diluted the strategy’s focus on the effectiveness and appropriateness of programs and reduced evaluation’s reference to the process of policy advising. It was argued that PMRs were the product of bureaucratic expedience which displaced initial concerns with public accountability and represented an attempt by central agencies to extend their influence over the processes of policy advising. In other words, PMRs were an attempt to provide leverage over that subset of the ‘rules’ of policy coordination relating to the structures and processes of IDCs. In this way the evaluation of policy advice is emerging as one instrument for enhancing central agency policy management. It would, however, strain the credibility of the available evidence to extrapolate any general and enduring change in policy development and coordination processes.

Question Two. To what extent has program evaluation altered the role of central agencies?

This second question was directed at establishing the extent to which the introduction of systemic program evaluation may have influenced the policy management roles of central agencies, and in particular the resource coordination dimensions of the central budget agency, the Department of Finance. There was some suggestion that the trend for central agencies, but most acutely for budget agencies, increasingly to involve themselves in questions of policy substance may be expediting the conditions for what was tentatively referred to as "policy overstretch". This was defined as the situation that a central agency faces when it takes on more policy coordination obligations than it can deal with adequately because it can never possess appropriate levels of policy expertise (see above, p45). The result may be impaired coordination capacities.

The dangers of policy overstretch in central budget agency policy management are always prevalent. The area of expertise for budget agencies, such as the Department of Finance, is advising on the financial implications of policy proposals. However, as suggested earlier, expenditure and policy are essentially two sides of the one coin. Commentary on the financial implications of a budgetary proposal must, in most if not all instances, also involve consideration of its substantive policy implications. When the Treasury Department also had carriage for expenditure review, the chief complaint from line departments was that department's insistence on challenging the "advisability, rather than the cost" of policy proposals (Weller & Cutt 1976, p49). This stance was perhaps more justifiable when the Treasury was the *de facto* policy coordinator in the absence of credible alternatives (such as DPMC).

For the Department of Finance, the evaluation strategy and its focus on the effectiveness of policy can be interpreted as a thinly disguised attempt to break the legacy of its institutional origins within the Department of the Treasury. The FMIP in general, and the evaluation strategy in particular, are excellent examples of how the Department of Finance approached the task of defining a "discrete and defensible role" within Australian government administration (Wanna, Kelly and Forster 1996, p4). Program evaluation's contribution to this definitional task was to secure either one or both of the following two objectives. First, the evaluation strategy could be viewed as a bid by the central budget agency to extend and augment its level of expertise in substantive policy areas, and transform the nature of budgetary choices from a purely financial perspective that aspired to a 'pseudo-policy' determination, to a genuine consideration of the substance of policy and its wider implications. Second, as discussed above, the evaluation strategy sought to alter the rules of the budgetary process in order to formally recognise the evaluation of policy effectiveness. In this it sought to establish a new common information base, sourced from line departments themselves, on which budgetary negotiations between the budget agency and departments would proceed. To this extent, program evaluation can be seen as an attempt by the Department of Finance to rationalise and hence legitimise the historically high levels of Treasury intervention into line department policy content as well as the review methods used to facilitate this (cf Weller & Cutt 1976, pp49-51, 78-84).

As discussed above, the evaluation strategy had only modest success in its quest to transform the norms and procedures of the budgetary process. In part this can be attributed to flaws in both policy design and implementation, particularly with respect to evaluation planning, but it also has to do with the fact that the expenditure review task facing budget officials has always been in the nature of *policy analysis*, since the financial and policy considerations of proposals are simply too difficult (or impractical) to separate. In effect, while the evaluation strategy may have increased moderately the awareness of program effectiveness in the work of budget officials, this was not ordinarily transferred to their advice on expenditure policy. At the same time, we could also conclude that program evaluation has not exacerbated the potential for policy overstretch. However, while financial and policy considerations are difficult to distinguish clearly, the budget agency is still confronted with a choice between maintaining a policy analysis capability that claims to deliver expert commentary on policy proposals or artificially confining their advice to the financial implications of policy proposals.[1] As we saw in Chapter Seven, the decision to intervene in an expert capacity, whether in budgetary decision making or through coordination mechanisms such as IDCs, is essentially a

[1] Of course, it is the definition of "financial implications" that constitutes the core of any concerns about the role of central budget agencies. The problem, as line departments would see it, is that the Department of Finance's definition of "financial implications" is likely to cast an exceptionally wide net.

matter of judgment. It is also a strategic choice complete with different pay offs. These are presented in **Figure 8.1** below.[2]

At a general level the choice for a central budget agency is likely to be between maintaining *specialist* capabilities in the financial aspects of expenditure and cultivating *generalist* expertise across all the policy responsibilities covered by the administration. In **Figure 8.1** this choice is represented on the horizontal axis. Specialist capacities are the distinctive contribution that the central budget agency brings to the policy process, while the generalist position duplicates policy capabilities already existing in other departments (on which see Uhr 1996 (forthcoming), pp18-21). At the same time, a central budget agency can choose to intervene on questions of either policy process or policy substance, represented by the vertical axis in **Figure 8.1.** Questions of process refer to the "rules of the game" that a central budget agency will maintain. These might include the procedural aspects of administrative or jurisdictional overlaps and the establishment and maintenance of mechanisms for resolving more substantive policy disputes. In the case of the Department of Finance this also involves maintaining the expenditure management system. On the other hand, substance refers more to the content of specific policies and advising on its merits.

Figure 8.1. Strategic Choices Facing Central Budget Agencies in Resource Coordination

	Specialist	Generalist
Process	1	3
Substance	2	4

[2] This table is derived from a framework developed by Professor Beryl Radin for analysing the different roles of central budget agencies (personal correspondence, August 1995).

From **Figure 8.1** the two extreme positions can be found in quadrants 1 and 4. In quadrant 1, the central budget agency has confined itself to a specialist role of advising on the financial aspects of proposals (narrowly defined) and maintaining the processes of expenditure management. Quadrant 4 is where the budget agency assumes a generalist policy advising capacity and duplicates (in a limited way) the expertise existing in most policy areas *and* advises on the merits of policy based on an assessment of the appropriateness of policy. In quadrant 4 the potential for policy overstretch is greatest, while in quadrant 1 the risk of a dissipating whole-of-government perspective is at its extreme. Under the evaluation strategy, assessment was supposed to be provided by program evaluation as a commonly accepted information base. Instead, evaluation has played at best a marginal role and budgetary negotiations (at least from the budget agency perspective) continue much as before.

As we have noted, budget officials already consider themselves to be generalist policy analysts. On this, program evaluation can not be charged with exacerbating any tendencies towards policy overstretch. Because of the symbiotic relationship between expenditure and policy, to use the distinction made by the Uhr Review between assessment and evaluation, discussed in the previous chapter, the task facing a budget agency such as the Department of Finance is to judge when the assessment (or specialist) role applies and when the evaluation (or generalist) role applies.[3] This is obviously a contingency based decision that could see different strategic choices (as outlined in **Figure 8.1**) applied for different coordination tasks. It is also heavily dependent on the professional expertise of budget officials. Evaluation in the form of PMRs, however, may hold some potential for marking out the limits to the generalist policy role contained in quadrant 4 and hence minimising the risk of policy overstretch.

To illustrate let us take a recent example of policy development where signs of policy overstretch may have arisen. *Plane Safe*, the report of the House of

[3] An alternative is to give a more policy balanced player increased leverage in the task of resource coordination—that is, give say DPMC an expanded role in vetting budget proposals and judging on the merits of policy proposals from a whole-of-government perspective (as envisaged by the *Report* of the RCAGA, see pp67-68 above). This would require a fundamental reorientation of the central machinery of government to shift the focus of the budget from expenditure to policy concerns. The evaluation strategy has failed to categorically achieve this from a coordination perspective (although it may well have had more success at the portfolio level).

Representatives Standing Committee on Transport, Communications and Infrastructure Inquiry into Aviation Safety was prompted by a series of fatal commuter aircraft accidents between 1993 and 1994, and examined failings in the aviation safety regulation and review functions performed by the then Civil Aviation Authority (CAA) (HRSCTCI 1995). During the course of its examination the committee came to the conclusion that any inadequacies of the aviation safety regime could be traced back to advice provided to the government in 1990 that reduced public funding for the setting and supervision of safety standards, and reorganised both the commercial and regulatory arms of the CAA on a full cost recovery basis (HRSCTCI 1995, pp4, 15-16; *The Canberra Times*. 25 June 1995. p11). The committee made an attempt to "examine the quality of the advice provided to decision makers at the time the decision was made". This course yielded an admission from the Department of Finance "that there were better ways of addressing the cost recovery issues" (HRSCTCI 1995, p4). Perhaps more significantly, the Finance representative conceded that those sections of central agencies overseeing aviation safety had insufficient expertise to detect failings—"particularly technical aviation safety failings"—and that in reviewing CAA programs there were concerns "to the extent that the system did not reflect the current experience" (*The Canberra Times*. 25 June 1995. p11). The committee noted that despite recognition by aviation experts of the emergence of safety problems in the CAA during the period after 1990, the central coordinating agencies—including Finance—either neglected, or were unable, to follow these up.

This appears to be a clear case of "policy overstretch"—the central agencies, and in particular the Department of Finance, having set the cost recovery framework for the sector, had insufficient policy expertise appropriately to advise on either the effectiveness or adequacy of aviation safety programs. The question this poses is to what extent could these failings have been averted if there was regular systematic monitoring of policy development processes through, say PMR-like instruments? Would these measures have picked up anomalies in the system of policy review which failed to adequately inform the advice generated by the department? Or do these kinds of issues go too deeply to the substance of policy advice, the acceptance of which is ultimately a matter for the responsible minister, and constitutes an area which PMRs are not designed to address? We should note, for example, that a good deal of advice relating to program funding issues provided at the end of the 1980s was framed around full cost recovery mechanisms (Keating 1994, p4). If policy advice was reflecting a 'general climate of opinion' then two implications emerge—only hindsight will cast the advice as inadequate in light of subsequent outcomes, or the policy work of public agencies require *evaluation* mechanisms linked to external points of reference, such as alternative advice sourced from the general and expert public, that oblige advisers to at least acknowledge and consider options which may not coincide with political convenience (Uhr & Di Francesco 1994, p264). Because the process of policy advising often determines the complexion of the substance of policy advice, the potential merit of evaluation mechanisms such as PMRs lie in their influence on policy development processes. At the same time, retrospective analysis of policy development and coordination processes may indicate those policy areas from which central agencies should withdraw so as to achieve a more appropriate balance between general coordinating interests (in this case the deregulation of the civil aviation sector) and policy specific expertise.

Political Choice And The Evaluation Strategy As Policy Design

After the consideration of the specific implications of the evaluation strategy for the policy management functions of central agencies, what remains is to examine the main conclusions that emerge from this study with respect to the design of program evaluation policy within government more generally. There are principally two related issues that arise here—the *location* of the program evaluation function within the administration, and the extent to which it is able to satisfy the different *information needs* of decision makers. Taken together these comprise the most basic issue of determining clear objectives for a program evaluation policy. Who or what does program evaluation serve? Can we say that the evaluation strategy resolved either of these issues?

Firstly, what does it serve? The interpretation offered here is that the evaluation strategy served decision making in central government. However, while this objective was promoted categorically by the Department of Finance, the design of the evaluation strategy was far from clear since it purported to serve a range of purposes—decision making in central government, public accountability purposes and departmental program management. This study has concentrated only on the first of these and it would be imprudent to comment too widely on the remaining aims. Nonetheless, although the evaluation strategy was admirably comprehensive in its scope, selling program evaluation as a multi-purpose tool has diminished the full benefits that effectiveness evaluation might have offered to any one of the competing purposes. If the primary purpose of evaluation is to inform central priority setting—in terms of either budgeting or more general policy coordination—then there must be an explicit recognition of this in any evaluation policy as well as an acknowledgment that this is essentially a *coordination* role.

This study has demonstrated that this role is highly dependent on the demand by central agencies for evaluation information (which, in the case of the Department of Finance supply divisions, was not overwhelming) and the capacity and willingness of those agencies to use it. It has also been suggested that, if directed at central government, some degree of centralisation is required, where the responsible central agency must play an enforcement role. In a situation where line departments are responsible for producing information which may be used against them in the allocation of resources, there is very little incentive to provide this information voluntarily, and little comfort in the knowledge that it was "owned" by the department. If, on the other hand, program evaluation is essentially a responsibility of line departments for internal departmental program management purposes, then departments should be left to their own devices in evaluating their programs and central agencies would be better served in keeping a safe distance. The kinds of problems observed in this study arise when the two incompatible objectives are pursued simultaneously.

This brings us to the second issue of satisfying information needs—who does program evaluation serve? Ministers, the bureaucracy, or the wider community? Ideally, as the Australian evaluation strategy sought, all three. Under the strategy, users comprised a disparate group that included ministers (both finance and spending), various levels of the bureaucracy, Parliament and its agents, as well as the general community. This was obviously a function of the three objectives referred to above, but given that the leading aim transpired as supporting decision making in central government, based on the findings of this study we would be hard pressed to conclude that any users were particularly well served. We know that evaluation was intentionally directed at different information needs—the distinction between AEP evaluations (that addressed efficiency and process issues) and PEP evaluations (that looked at more strategic questions of program effectiveness and appropriateness) is the best example—and produced qualitatively different types of information. We also know that at the central agency level the kind of effectiveness information that might have assisted priority setting was not routinely available or, when it was, inconsistently relied upon. In part, this is reflected in the shift away from program evaluation to purpose specific program reviews in resource coordination. Information needs at the central level, whether central agency or ministerial, were not being served as anticipated. (Utilisation at the departmental level is a separate issue that requires investigation.)

The confusion over policy objectives underlines another unresolved, or perhaps simply neglected, question as to whether the political executive actually desire the effectiveness of their programs to be evaluated. Speaking of program evaluation policy in Canadian federal government in the early 1980s, but surely having wider application, one senior public official warned that

> . . . we must be sure that we do not try to press on politicians a system which asks the deeper questions of program *raison d'etre*, unless politicians really want to examine them at that level. . . An evaluation system designed by and for the bureaucracy will not work for politicians (Hicks 1981, p356).

This last observation is quite prophetic for the evaluation strategy. As we have seen, the long gestation period of the evaluation strategy was punctuated by ministerial and, much more commonly, extra-government parliamentary demands for evaluation to provide information on the performance and objectives of government policy. This parliamentary interest was motivated by issues of both public accountability and program management that was more responsive to community needs. Later, the evaluation strategy was to have explicit ministerial support. Overall, however, and especially with the acceleration of public sector restructuring in the mid to late 1980s, ministerial involvement was minimal and the control of the evaluation agenda was left in the hands of the Department of Finance. Effectiveness reviews were long seen as a mechanism for increasing ministerial control of policy making in central government, but while in form the evaluation strategy may have aspired to this goal, in substance it fell away in meeting the information needs of either ministers or the central agencies.

In this, the experience of the evaluation strategy reaffirms the importance of *political choice* in government policy making. Here the term "politics" is used in a specific sense, adapted from the framework developed by Weiss for analysing the basis of public policy positions (Weiss 1983). This takes as its starting point the acknowledgment that, because public policy is usually centred on the capacity of governments to accommodate competing societal interests, political choice must always be a legitimate characteristic of decision making in government:

> Observers who expect the subcategory of information that is social science research to have immediate and independent power in the policy process, and who bitterly complain about the intrusion of "politics" (i.e. interests and ideologies) into the use of research, implicitly hold a distorted view of how decisions are made (Weiss 1983, pp221-222).

It is the fluid interaction between information (which, for this study of program evaluation, was equated in Chapter One with analysis) and politics that determines how and when information is likely to have influence in decision making. "Politics" can, in parallel with Weiss, be taken to include both ideology (or the value positions and political orientation that decision makers may subscribe to) as well as that stalwart of political analysis: self-interest, broadly defined. Although in these terms politics encompasses predispositions on the part of decision makers, these are by no means always fixed, but instead subject to continuous rounds of redefinition in accordance with the specific situation faced by decision makers:

> the interaction is constant and iterative, and policy makers work out the specification of their ideologies and interests *in conjunction with* their processing of information (Weiss 1983, p229; emphasis in original; see generally pp223-237).

Hence, if a source of information seeks to affect (or, more succinctly, reduce) the "politics" of decision making, then it is likely pursuing what could be seen as an unrealisable goal, since it is effectively denying the legitimacy of "politics" in those same processes. Information is properly only one of a number of competing factors that may influence decision making in government. For Weiss, the often observed outcomes of this interaction are gradual movements in policy punctuated by change forced by shifts in power—such as with a change in government—that alter markedly the balance between competing sources of ideologies, interests and, ultimately, information (Weiss 1983, pp239-241). Information may have either more or less influence.

Looking to the political dynamics of decision making is one way to rationalise the bounded role of program evaluation, but there are limits to its explanatory power. Pointedly, it suggests that policy stability is the rule, and policy change the exception. Conversely, it can also be argued that this rule is the exception, and that in order to have influence program evaluation (or any analytical technique directed at the effectiveness and relevance of policy) requires a period of policy stability. As some students of the macro policy process have suggested, in most pluralist political systems government control of policy and its outcomes is particularly weak precisely because goals are "easily reversible", so that evaluation policies themselves commonly are not installed or, if they are, prevented from working through neglect (Dunsire 1991, pp335-336; Stewart 1992, pp252-253). In these circumstances,

> the price of political reversibility may be taken in lack of policy evaluation. Moreover, for many political purposes, the trumpeted designation of the selected policy objective *is* the pay-off, and subsequent performance (implementation, 'success') in respect of that policy is either too remote a consideration to be of present interest, or it is positively unwelcome. Few politicians like to allow doubt that their policy is the right one, or to be told that it has not worked. Where the objectives of public policy are the legitimate prizes of politics, and the fairly rapid alternation in office of political parties with opposed value-systems is regarded as a *virtue*, then there is simply not a big enough premium on policy evaluation and correction by feedback mechanisms (Dunsire 1991, p335, emphasis in original).

If this was to be converted into the terminology employed by Weiss, it would be rare indeed for the balance between the ideology and interests of politicians to allow for the production and consideration of evaluation information—system dynamics, apparently, would not permit it.

The experience of the Australian Labor Government, by these standards, is exceptional. Despite the economic and social policy upheaval of the 1980s and the first half of the 1990s, the Labor Government held office continuously for five terms and provided a good degree of policy stability. As a consequence, even though for all intents and purposes it transpired to be a Department of Finance initiative, the evaluation strategy represented a credible attempt at installing a *system* of program evaluation. It exploited the "policy corridor" provided by two developments; first, the development of evaluation policy that had accumulated since the RCAGA, and second, the spur provided by the acute economic crisis of the mid 1980s. Even though it appears to have effected only modest levels of change, program evaluation was intended to serve distinct purposes during the late 1980s and early 1990s in terms of both modulating aggregate fiscal contraction *and* imposing fiscal discipline on individual departments.

However, when policy stability comes to an end, and the balance between ideology and interests is altered, the restricted reach of program evaluation is accentuated. Political regime change introduces relatively unfettered political choice. The most instructive example of this has been the change of government at the March 1996 federal election. The new Liberal-National Coalition Government came to power with a resounding lower house majority and a relatively innocuous agenda to maintain policy course where existing policy had worked. Policy change, except for those areas which were explicitly outlined in the party platform (such as industrial relations reform and the partial privatisation of the publicly owned telecommunications corporation, Telstra) would be marginal. This agenda, however, was quickly overrun in the wake of a projected underlying budget deficit for 1996-97 of some $8 billion (*The Australian.* 13 March 1996. pp1, 4-5). Immediate policy priorities turned to swift deficit reduction and 'cutback management' within a context of the government's election commitment not to resort to revenue (tax raising) measures. Large expenditure reductions were foreshadowed and justified on the basis of the overwhelming "mandate" received by the new government. Therefore, the most urgent task facing the government was to determine policy priorities and excise or diminish program expenditure which was considered to be ineffective. The fascinating question is, on what information base would and could decisions be made?

Since the fiscal rectitude of the new Coalition Government has been even more pronounced than either of the second and third Labor Governments, it has been an almost natural reaction to further centralise decision making. The Howard

Government has certainly maintained (if not intensified) the primacy of the central economic agencies, Treasury and Finance, in driving policy choices (*The Australian Financial Review*. 8 August 1996. pp1, 16; *The Canberra Times*. 8 August 1996. p13; 10 August 1996. p11). This reliance could be evidenced by the involvement of these agencies in supporting the agenda setting role of the Coalition Government's National Commission of Audit, and has grown stronger despite a climate of drastic pruning in public service resources. As a consequence, the 1996-97 Budget was an excellent test for the endurance of the evaluation strategy under the new Coalition government. In particular, guided by its commitment to deficit reduction, and compelled by the growing number of damaging leaks from the bureaucracy, the government adopted a strategy of announcing savings measures in advance of the Budget in order to minimise opposition at the time of its release. From this it is possible to examine the rationale for expenditure policy in a number of areas targeted for change.

As could be expected, in most if not all cases, the government publicly justified decisions to cut program expenditures on the basis of either the inefficiency or ineffectiveness of programs or, less commonly, with reference to the inappropriateness of a program's objectives. Some of the more prominent examples of the latter include the abolition of regional development programs and the 10 per cent reduction in recurrent funding to the Australian Broadcasting Corporation (ABC), both of which were seen to be inefficient *and* inappropriate (or unfit) given the changed policy priorities of the government (in the case of the ABC the responsible Minister claiming that the broadcaster unduly emphasised certain social issues such as "racism and aboriginal affairs" at the expense of other mainstream issues, such as economics (*The Sydney Morning Herald*. 18 July 1996. p8)). There has been almost no attempt publicly to justify decisions based on performance information already available within the bureaucracy. In the two cases where an evaluation has been invoked, the information has been employed selectively to smooth the passage of policy decisions already made.

The first example relates to the streamlining of labour market programs within the Employment, Education, Training and Youth Affairs portfolio (*Australian Financial Review*. 2 August 1996. p4; *The Australian*. 2 August 1996. p1; Media Release from Senator Vanstone, "Report Criticises Working Nation", 1 August 1996). The responsible minister had indicated that the limited funding available because of deficit reduction measures would have to be directed at the "most efficient programs" and "those programs that generate the best results for job seekers"—that is, resources would be allocated on the basis of program effectiveness (Vanstone 1996a, pp9-10). As a consequence, Cabinet decided to "streamline" the existing Working Nation labour market programs in order to both produce some $1.3 billion in savings and redirect funding to its Modernised Apprenticeship and Traineeship System (MATS). The decision coincided with, and "received some support" from, the release of a DEET evaluation of Working Nation programs. The evaluation contained ambiguous findings. On the one hand, the level of long term unemployed had decreased significantly, but on the other it was found that case management was not resulting in "unsubsidised employment outcomes for a large proportion of disadvantaged clients"—the reductions in unemployment rates could, according to the Minister, be attributed to the reclassification of the long term unemployed as short term unemployed after their completion of a program placement. The evaluation had "showed Working Nation to be a cruel hoax and showed the fall in the long-term unemployed statistics was clearly misleading". In this case an evaluation was used selectively to justify policy change.

The second case relates to the scrapping of the 150 per cent tax concessions for syndicated companies engaged in research and development. The decision, which took effect well before the August 20 Budget, *coincided* with the release of a report by the (then) Bureau of Industry Economics (BIE) which recommended the abolition of the program (*The Sydney Morning Herald*. 24 July 1996. p1; *The Australian Financial Review*. 24 July 1996. p7). The BIE report argued that:

> In some, possibly many, individual syndicates, the Government is providing subsidy rates for R&D in excess of 100 per cent and producing sizeable negative returns to Australians . . . We are no longer certain that the program generates net benefits for Australians (*The Australian Financial Review*. 24 July 1996. p7).

The BIE report clearly points to the ineffectiveness of the program—it was simply not delivering value for money for Australian taxpayers. According to the Treasurer there were a number of reasons underpinning the decision. Ineffectiveness of the BIE variety was one—the system "invited investors to rort it" and "attempts to try to limit it are always behind the game: the only way of dealing with the problem is to remove eligibility for syndication altogether". On the other hand the termination was also a part of the government's attack on "business welfare", and under these circumstances "we [the government] couldn't say to the battler with a straight face that nothing should be done about it" (*The Sydney Morning Herald*. 24 July 1996. p1). The tenor of these comments suggest that the BIE report was employed strategically to publicly justify a decision that had already been taken.

Of course, this can all be explained quite simply as a case of changed political priorities. As such, the imperatives of political choice will outweigh analytical information because the objectives have changed. Using the previous government's evaluation strategy the new Coalition Government could plausibly argue that the appropriateness criterion applies—that is, government priorities have changed and therefore any evaluative information framed around superseded objectives is invalidated. This would certainly be consistent with the policy relevance thrust of the evaluation strategy. In all likelihood this *was* the position adopted by the new government. Alternatively, there are good reasons to believe that the Department of Finance had already been targeting many of the program cuts that would frame the contraction agenda of the new government. The Department's search for program savings, especially those relating to the ABC, and the industry, health, and higher education portfolios, had indeed commenced under the previous Labor Government (*The Sydney Morning Herald*. 16 March 1996. p1, 8). If one were to extrapolate from the findings relating to the role of program evaluation in the budgetary process (Chapter Six), it could be said that these recommendations for savings measures were developed with limited assistance from effectiveness or appropriateness information.

CONCLUSION AND PROSPECTS

To recapitulate, the thesis has made the case for two main arguments. The first, based on a detailed historical account of the development of program evaluation policy, has been that the evaluation strategy was designed to serve central government decision making and hence the policy management functions of central agencies. Despite being the product of budgetary stress, the evaluation strategy was distinguished by its attempt to promote substantive policy ahead of purely fiscal concerns in the priority setting of central agencies. The second argument, grounded in an analysis of program evaluation's contribution to two aspects of central agency policy management, suggested that the evaluation strategy had fallen short of achieving the desired level of enduring change in those functions. One final task might be to offer some observations on the prospects for program evaluation in Australian government, taking account both of the impacts noted in the study and emerging trends in the structure of government in Australia. What follows is conjectural, but may provide some leads for further research in this area.

One area that is ripe for further investigation remains the residual objectives of the evaluation strategy. This study confined itself to the application of program evaluation to the policy management functions of central agencies, but what of program evaluation *within* line departments and agencies? We have seen that the stewardship role of the Department of Finance for the evaluation strategy did not extend to the conduct, use, cost and impact of program evaluation in departments for their own program management and policy development purposes. Certainly, some illumination is provided in the series of reports undertaken by the Audit Office, but these are duly restricted to the management of procedures for the conduct of program evaluation (ANAO 1991a; 1992a; 1993). With this in mind, some of the questions that deserve consideration might include the following. To what extent are evaluation practices an established part of day to day program management? Is the evaluation undertaken in portfolios in line with the models promulgated by the Department of Finance? If not, why? Has evaluation strengthened the linkages between results management and the needs of program clients? Who conducts the evaluation? What proportion of evaluation activity is undertaken by outside consultants? Is there a perceived difference in the utility of this evaluation? What is the level of portfolio ministerial or senior executive demand for program evaluation?

Another issue that lingers is program evaluation for public accountability. As we have discussed, the responsibility for the effectiveness and appropriateness of government programs remains securely located with Ministers—there appears to have been no appreciable increase in the responsibility of the public service for these matters. Indeed, the final phase of the evaluation strategy which sought to evaluate policy advising programs marked a conspicuous return to the safety of the policy-administration dichotomy. In retrospect it appears as though the RCAGA hit the mark in categorising program evaluation as a poor mechanism for public accountability purposes. This is not only because the strategy itself was primarily designed not to satisfy these obligations, but also because the bounds of public accountability in terms of policy effectiveness and appropriateness remained unsatisfactorily ill-defined and the reach of parliament into the executive severely circumscribed. In the words of one former parliamentary officer:

> In the context of the accountability of the executive to parliament, there is a lack of fit between the information provided by evaluation, and the capacity and will to use it or to have recourse against the subjects of the evaluations (Hamburger 1992, p79).

Just as parliament played a key role in pushing for the adoption of effectiveness review within the executive during the late 1970s and early 1980s, so too should it give further consideration to developing mechanisms for not only reviewing performance information derived from evaluations but also for enforcing any remedial action. The one avenue that remains would be to wrench program evaluation from the executive, recast it as an independent function akin to financial statement and performance auditing, and locate it with the Audit Office. The role of program evaluation would be to assess the effectiveness of government policy against commonly defined criteria and report the results to a standing committee of the parliament which would, in turn, require the responsible minister and department to explain any deficiencies and take corrective action within a specified period of time. Such a proposal perhaps borders on the fanciful, since it presumes both parliamentary will and wherewithal, but ultimately, if program evaluation is to serve public accountability purposes, it must be reforged *out* of executive and *into* parliamentary terms.

Finally, what are the implications of changes in the scope and role of the public sector for a program evaluation policy? The change of government in March 1996 has certainly placed the continued public provision of a number of government services on notice, by canvassing further restructuring of the public service in favour of a contract-based framework for the delivery of public goods and services. The Coalition Government’s National Commission of Audit, itself a form of evaluation, has foreshadowed a range of reforms centred on restructuring the delivery of health and welfare services around purchaser/provider roles (where the government purchases services for specified clients from third party providers) as well as the application of market contestability principles (NCA 1996, pp13-17, 54-55, 101-104). Prominent examples of recommended restructuring along purchaser-provider lines include completing the contracting out of the delivery of services common to government administration, and the placement of unemployment case management and labour market services on a highly competitive footing through the creation of "employment placement markets" (on which see Vanstone 1996b, pp17-25). If the Australian public sector is moving increasingly towards the "contract state", what is the role for program evaluation which, in the absence of the price mechanism, was intended to provide a proxy measure of value for money in public sector operations?

There are a number of possibilities. Evaluation might become an important tool for the regulation and monitoring of contractual performance by non-government service providers. In this case, the public service would be pruned back to a core policy development, contract specification and monitoring role, requiring some sort of instrument to assess contract performance in terms of policy effectiveness and compliance. Evaluation could serve this role. This does, however, have the unsatisfactory side-effect of introducing complications for public accountability (see, for instance, the New Zealand experience in Martin 1995; Gregory 1995). With the increased fragmentation between ministers, public servants and contractors of responsibility for policy effectiveness, the question of who and what evaluation would serve are exacerbated. It could serve an accountability role for the political executive, but it may reduce the receptivity of contractors if it publicly discloses the deficiencies of the contractor; it could be used by contractors to increase program responsiveness, but this would likely preclude accountability purposes on grounds of commercial confidence; it could be used by parliament as a review mechanism for ensuring that program delivery is reflecting adequately community demands. Alternatively, program evaluation might be jettisoned as unworkable or immaterial.

But this represents only a possible future. The analysis of the current evaluation strategy undertaken in this thesis suggests that the role of effectiveness review within Australian government has been a captive of its own policy history. Despite its aspirations to achieving "evaluation for all seasons", the evaluation strategy has been intentionally framed around the coordination questions facing central agency policy management, particularly those relating to budgetary decision making. In the case of resource coordination, program evaluation has, on the evidence presented, fallen considerably short of fundamentally recasting the budgetary process along policy effectiveness lines. At the same time, the extension of the evaluation strategy to policy advising programs has witnessed a retreat from the strict evaluation of outcomes and, in doing so, effectively re-marked the limits to public service responsibility for policy. Institutionalising policy effectiveness as a review criterion of and for the executive has proved to be a difficult and fatiguing task; and, ultimately, less like tilling fertile soil and more like ploughing the policy sands.

Bibliography

Table of Contents

Primary Sources

Secondary Sources

Primary Sources

(A) Archival Material

Australian Archives Series Number A 5433. Files on the Budget Estimates Inquiry. Item 'BE' single number series (1-16). House of Representatives Standing Committee on Expenditure.

Australian Archives Series Number AA 1982/568. Papers on the Reference on Health and Welfare Evaluation 1976-1979. Senate Standing Committee on Social Welfare.

(B) Documents Of The Commonwealth Department Of Finance

A number of selected files and documents were made available to the author by the Department of Finance in the period March to July 1995. They are cited throughout the thesis as (FD: document number).

(C) Parliamentary Debates

Commonwealth Parliamentary Debates (CPD). House of Representatives.

8 April 1976; 18 November 1976; 7 November 1977; 12 September 1978; 25 October 1978; 3 April 1979; 4 March 1980; 25 September 1986; 15 September 1987; 17 September 1987.

Commonwealth Parliamentary Debates (CPD). Senate.

6 November 1979; 23 May 1980.

Commonwealth Parliamentary Debates (CPD). Senate. Estimates Committees.

Committee D. 28 April 1977, p61
Committee C. 2 May 1978, p49
Committee A. 25 September 1990, A313-A317
Committee A. 9 September 1991, A125-A130
Committee A. 25 September 1991, A463-A464
Committee A. 25 September 1991, A545-A546
Committee D. 2 September 1993, D249

(D) Government and Parliamentary Publications

Amies, M. 1995. *Evaluation in the HSH Portfolio: Adding Value to Policy Development and Program Performance.* 23 May. Portfolio Strategies Group—Evaluation Directorate. Canberra.

ANAO (Australian National Audit Office). 1991a. *Implementation of Program Evaluation - Stage 1.* Audit Report No. 23 1990-91. AGPS. Canberra.

ANAO. 1991b. *Evaluation in Preparation of the Budget.* Audit Report No. 13 1991-92. AGPS. Canberra.

ANAO. 1992a. *Program Evaluation in the Departments of Social Security and Primary Industries and Energy.* Audit Report No. 26 1991-92. AGPS. Canberra.

ANAO. 1992b. *Department of the Treasury: Procedures for Managing the Economic Policy Program.* Audit Report No. 36 1991-92.. AGPS. Canberra.

ANAO. 1993. *Program Evaluation: Strategies, Practices and Impacts. Industry, Technology and Regional Development Portfolio.* Audit Report No. 35 1992-93. AGPS. Canberra.

Argy, F. 1993. *Review of Treasury Policy Advice on the Government Response to the Martin Inquiry (A Pocketful of Change).* A confidential report prepared for the Treasury Department. May. Department of the Treasury. Canberra.

Canada. 1962. Royal Commission on Government Organization. *Management of the Public Service—Volume 1.* Queen's Printer. Ottawa.

Cmnd. 3638. 1968. *The Civil Service. Vol. 1: Report of the Committee 1966-68.* (Chairman: Lord Fulton). HMSO. London.

Commonwealth of Australia. 1977. *Commonwealth Government Directory*. AGPS. Canberra.

Commonwealth of Australia. 1978. *Commonwealth Government Directory*. AGPS. Canberra.

Commonwealth of Australia. 1983. *Reforming the Australian Public Service, A Statement of the Government's Intentions.* December. AGPS. Canberra.

Commonwealth of Australia. 1984. *Budget Reform—A Statement of the Government's Achievements and Intentions in Reforming the Australian Government Financial Administration.* AGPS. Canberra.

Commonwealth of Australia. 1993a. *Program Performance Statements 1993-94—Finance Portfolio.* Budget Related Paper No. 7.6. AGPS. Canberra.

Commonwealth of Australia. 1993b. *Budget Statements 1993-94.* Budget Related Paper No. 1. AGPS. Canberra.

Commonwealth of Australia. 1994. *Budget Statements 1994-95*. Budget Related Paper No. 1. AGPS. Canberra.

Commonwealth of Australia. 1995. *Budget Statements 1995-96*. Budget Related Paper No. 1. AGPS. Canberra.

DEET (Department of Employment, Education and Training). 1994. *Review of the Evaluation Function in DEET*. February. AGPS. Canberra.

DFAT (Department of Foreign Affairs and Trade). 1991. *Submission by the Department of Foreign Affairs and Trade to the Inquiry into the Management and Operations of the Department of Foreign Affairs and Trade by the Senate Standing Committee on Finance and Public Administration*. November. Department of Foreign Affairs and Trade. Canberra.

DHSH (Department of Human Services and Health). 1994. *Portfolio Evaluation Plan 1995-97*. AGPS. Canberra.

DHSH. 1995a. *Portfolio Evaluation Plan 1996-98: Strategies for Performance*. November. AGPS. Canberra.

DHSH. 1995b. *1994-95 Annual Report*. AGPS. Canberra.

DoF (Department of Finance). 1979. *Annual Report 1978-79*. AGPS. Canberra.

DoF. 1985. *Annual Report 1984-85*. AGPS. Canberra.

DoF. 1986a. *Annual Report 1985-86*. AGPS. Canberra.

DoF. 1986b. *Program Budgeting and Policy Areas*. Program Budgeting Issues No. 11. July. Department of Finance. Canberra.

DoF. 1987a. *Annual Report 1986-87*. AGPS. Canberra.

DoF. 1987b. *FMIP and Program Budgeting: A Study of Implementation in Selected Agencies*. August. AGPS. Canberra.

DoF. 1987c. "Implementation of Evaluation Plans". Memorandum by B. Corcoran to Heads of Corporate Services in Departments and Relevant Authorities, 25 September.

DoF. 1988a. *Annual Report 1987-88*. AGPS. Canberra.

DoF. 1988b. *FMIP Report*. December. AGPS. Canberra.

DoF. 1989a. *Annual Report 1988-89*. AGPS. Canberra.

DoF. 1989b. *Program Evaluation: A Guide for Program Managers*. March. Department of Finance. Canberra.

DoF. 1989c. *Managing Evaluation within Departments and Agencies*. July. Department of Finance. Canberra.

DoF. 1990a. *Annual Report 1989-90.* AGPS. Canberra.

DoF. 1990b. *Submission by the Department of Finance to the Joint Committee of Public Accounts Inquiry into the Department of Finance.* 31 January. Department of Finance. Canberra.

DoF. 1991a. *Annual Report 1990-91.* AGPS. Canberra.

DoF. 1991b. *A Guide for the Preparation of Evaluation Reports.* November. Department of Finance. Canberra.

DoF. 1991c. *Promoting Value for Money: The Role of the Department of Finance.* A Summary of the Department of Finance Submission to the Joint Committee of Public Accounts Inquiry into the Department of Finance. April. AGPS. Canberra.

DoF. 1992a. *Evaluation Papers: Introduction to the Series and Overview.* Department of Finance. Canberra.

DoF. 1992b. *Commonwealth Financial Management Handbook.* AGPS. Canberra.

DoF. 1992c. *The Role of Evaluation in Decision-Making in the 1991-92 Budget Context.* Discussion Paper. June. Department of Finance. Canberra.

DoF. 1992d. *Keeping the Score: An Evaluation of Estimates Processes and Practices Within the Commonwealth Department of Finance.* June. Department of Finance. Canberra.

DoF. 1993a. *Annual Report 1992-93.* AGPS. Canberra.

DoF 1993b. *The Use of Evaluation in the 1992-93 Budget.* Discussion Paper. July. Department of Finance. Canberra.

DoF. 1994a. *Doing Evaluations: A practical guide.* Department of Finance. Canberra.

DoF. 1994b. *The Use of Evaluation in the 1993-94 Budget.* Discussion Paper. April. Department of Finance. Canberra.

DoF. 1994c. *The Use of Evaluation in the 1994-95 Budget.* Discussion Paper. October. Department of Finance. Canberra.

DoF. 1995. *Review of the Booklet 'The Use of Evaluation in the Budget'.* Draft. July. Department of Finance. Canberra.

DoF & DPMC (Department of the Prime Minister and Cabinet). 1989. *Towards A Fairer Australia. Social Justice and Program Management: A Guide.* AGPS. Canberra.

DPMC (Department of the Prime Minister and Cabinet). 1979. *Annual Report 1978-79.* AGPS. Canberra.

DPMC. 1984. *Annual Report 1983-84.* AGPS. Canberra.

DPMC. 1987. *Annual Report 1986-87.* AGPS. Canberra.

DPMC. 1989. *Evaluation of Policy Advising Programs: Discussion Paper.* April. Department of the Prime Minister and Cabinet. Canberra.

DPMC. Planning and Evaluation Branch. 1991. *Issues Paper: Performance Assessment of Policy Work.* A paper prepared by the Department of the Prime Minister and Cabinet for the Working Group on Performance Assessment of Policy Work. September. Department of the Prime Minister and Cabinet. Canberra.

Evaluation Task Force. Department of Finance. 1988. "Evaluation—Recent Developments in the Commonwealth Public Service". Paper presented to the National Evaluation Conference. Sydney.

Glenn, G. 1993. *Report on the Evaluation of the Policy Development of the 1992-93 Carers Package.* April. Department of the Prime Minister and Cabinet. Canberra.

Glenn, G. 1994. *Evaluation of the Forward Estimates Strategy Papers for the 1993-94 Budget.* April. Department of Finance. Canberra.

HM Treasury. 1988. *Policy Evaluation: A Guide For Managers.* HMSO. London.

HRSCE (House of Representatives Standing Committee on Expenditure). 1977a. *Accommodation For Married Servicemen.* May. AGPS. Canberra.

HRSCE. 1977b. *A Year's Experience.* October. AGPS. Canberra.

HRSCE. 1978. *Northern Territory Forestry Program.* May. AGPS. Canberra.

HRSCE. 1979. *Parliament and Public Expenditure.* February. AGPS. Canberra.

HRSCFPA (House of Representatives Standing Committee on Finance and Public Administration). 1990. *Not Dollars Alone: Review of the Financial Management Improvement Program.* September. AGPS. Canberra.

HRSCTCI (House of Representatives Standing Committee on Transport, Communications and Infrastructure). 1995. *Plane Safe. Inquiry into Aviation Safety: The Commuter and General Aviation Sectors.* December. AGPS. Canberra.

JCPA (Joint Committee of Public Accounts). 1979. *Financial Administration—Parliamentary Scrutiny.* Parliamentary Seminar. 1 June 1979. AGPS. Canberra.

JCPA. 1981. *Government Expenditure and Accountability—The Relationship Between the Parliament and the Public Service in the 1980s.* Parliamentary Seminar. 16 May 1980. AGPS. Canberra.

JCPA. 1982. *The Selection and Development of Senior Managers in the Commonwealth Public Service.* Report 202. AGPS. Canberra.

JCPA. 1989. *The Auditor General: Ally of the People and Parliament. Reform of the Australian Audit Office.* Report 296. AGPS. Canberra.

MAB-MIAC (Management Advisory Board-Management Improvement Advisory Committee). 1991. *Accountability in the Commonwealth Public Sector: An Exposure Draft.* No. 5. June. AGPS. Canberra.

MAB-MIAC . Task Force on Management Improvement. 1993a. *The Australian Public Service Reformed: An evaluation of a decade of management reform*. December. 1992. AGPS. Canberra.

MAB-MIAC. Task Force on Management Improvement. 1993b. *Agency Survey Responses*. February. TFMI. Canberra.

MAB-MIAC. 1993c. *Accountability in the Commonwealth Public Sector*. No. 11. June. AGPS. Canberra.

NCA (National Commission of Audit). 1996. *Report to the Commonwealth Government*. June. AGPS. Canberra.

NSWPSB (New South Wales Public Service Board). 1985. *Report of the New South Wales Public Service Board for the Year Ended 30 June 1985*. Government Printer. Sydney.

PSB (Public Service Board). 1975. *Public Service Board Memorandum No. 22 to the Royal Commission on Australian Government Administration*. September. Public Service Board. Canberra.

PSB. 1976. *Report for 1975-76*. AGPS. Canberra.

PSB. 1977. *Annual Report 1976-77*. AGPS. Canberra.

PSB. 1978. *Annual Report 1977-78*. AGPS. Canberra.

PSB. 1985. *Annual Report 1984-85*. AGPS. Canberra.

PSB & DoF (Department of Finance). 1984. *Financial Management Improvement Program Diagnostic Study*. AGPS. Canberra.

PSB & DoF. 1986. *FMIP Report*. June. AGPS. Canberra.

PSB & DoF. 1987. *Evaluating Government Programs: A Handbook*. May. AGPS. Canberra.

PSC (Public Service Commission). 1995. *Performance Assessment of Policy Work: Report of a Working Group*. Occasional Paper No. 1. September. AGPS. Canberra.

RCA (Review of Commonwealth Administration). 1983. *Report*. AGPS. Canberra. (Chairman: J. Reid)

RCAGA (Royal Commission on Australian Government Administration). 1976. *Report*. AGPS. Canberra. (Chairman: H. C. Coombs)

RNSWGA (Review of New South Wales Government Administration). 1977. *Directions for Change. Interim Report—November 1977*. Government Printer. Sydney. (Chairman: P. Wilenski)

SSCFPA (Senate Standing Committee on Finance and Public Administration). 1989. *Review of the Efficiency Scrutiny Program*. April. AGPS. Canberra.

SSCFPA. 1994. *Public Service Reform. Volume 2: Conference Proceedings*. August. Department of the Senate. Canberra.

SSCSW (Senate Standing Committee on Social Welfare). 1979. *Through a Glass, Darkly: Evaluation in Australian Health and Welfare Services.* Volume 1: Report and Appendixes. AGPS. Canberra.

Treasury (Department of the Treasury). 1991. *Annual Report 1990-91.* AGPS. Canberra.

Treasury. 1992. *Annual Report 1991-92.* AGPS. Canberra.

Treasury. 1993. *Annual Report 1992-93.* AGPS. Canberra.

Treasury. 1995. *Annual Report 1994-95.* AGPS. Canberra.

Uhr, J. 1995. *Review of Finance's Role in Promoting and Using Evaluation.* June. Department of Finance. Canberra.

Vanstone, A. 1996a. "Media Release on Labour Market Programs". 28 June 1996. *Ministerial Document Service*. No.1 / 96-97. 1 July 1996. pp9-10.

Vanstone, A. 1996b. *Reforming Employment Assistance: Helping Australians into Real Jobs.* Ministerial Statement by Senator the Honourable Amanda Vanstone, Minister for Employment, Education, Training and Youth Affairs. August. AGPS. Canberra.

Weller, P. 1995. *Review of Commonwealth-State Reform Processes.* June. Department of the Prime Minister and Cabinet. Canberra.

(E) NEWSPAPERS

The Australian

Gordon, M. & Henderson, I. "Costello pledges $8 billion in cuts". *The Australian*. 13 March 1996. pp1, 4-5.

Henderson, I. "Job schemes face $1.3 bn cut". *The Australian*. 2 August 1996. p1

The Australian Financial Review

Ellis, S. & Shires, D. "Business furious as R&D tax breaks slashed". *The Australian Financial Review*. 24 July 1996. pp1, 7.

Grattan, M. "Working Nation hoax blamed for jobless cuts". *The Australian Financial Review*. 2 August 1996. p4.

Grattan, M. "Treasury's New Life Under Howard". *The Australian Financial Review*. 8 August 1996. pp1, 16.

The Canberra Times

Henderson, I. “Treasury Department ignored, Cole claims”. *The Canberra Times.* 9 May 1995. pp1-2.

Taylor, M. “Culprit all around you, Mr Morris”. *The Canberra Times.* 25 June 1995. p11.

Wright, L. “Balance of Power Still the Name of the Game”. *The Canberra Times.* 8 August 1996. p13

Schroder, P. “Howard Painting Himself into a Corner on PS Cuts”. *The Canberra Times.* 10 August. 1996. p11

Brough, J. “Guard changes, ‘Godfather’ goes”. *The Canberra Times.* 18 June 1995. p10.

The Sydney Morning Herald

Cleary, P. “Razor gang gets $20 bn hit list”. *The Sydney Morning Herald.* 16 March 1996. p1

Cleary, P. “Wielding a razor is not so easy”. *The Sydney Morning Herald.* 16 March 1996. p8

Brough, J. & Passey, D. “Alston attacked over black bias claims”. *The Sydney Morning Herald.* 18 July 1996. p8

Stott, D. “$1 bn axed in research tax breaks”. *The Sydney Morning Herald.* 24 July 1996. p1.

Secondary Sources

(E) BOOKS AND ARTICLES

Alaba, R. 1994. *Inside Bureaucratic Power: The Wilenski Review of New South Wales Government.* Hale & Iremonger. Sydney.

Albaek, E. 1995. “Between Knowledge and Power: Utilization of Social Science in Public Policy Making”. *Policy Sciences.* 28 (1). pp79-100.

ALP (Australian Labor Party). 1983. *Labor and Quality of Government.* Policy presentation by Bob Hawke, Leader of the ALP and Gareth Evans, Shadow

Attorney-General and Chairman, Federal Parliamentary Labor Party Task Force on Government Administration. 9 February.

Amies, M. 1994. “Program Evaluation: A Commonwealth Perspective—Where Are We Now?”. *Evaluation Journal of Australasia.* 6 (1). pp31-42.

Aucoin, P. & Bakvis, H. 1993. “Consolidating Cabinet Portfolios: Australian Lessons for Canada”. *Canadian Public Administration.* 36 (3) Fall. pp392-420.

Aucoin, P. 1990. “Administrative Reform in Public Management: Paradigms, Principles, Paradoxes and Pendulums”. *Governance—An International Journal of Policy and Administration.* 3 (2) April. pp115-137.

Barrett, P. 1988. “Emerging Management and Budgetary Issues: The View From the Centre”. *Canberra Bulletin of Public Administration.* No. 54. May. pp52-57.

Barrett, P. 1992. “Evaluation as a Strategic Element of Reform in the Commonwealth Public Sector” in C. O’Faircheallaigh & B. Ryan (eds), *Program Evaluation and Performance Monitoring: An Australian Perspective.* Macmillan. South Melbourne. pp1-18.

Barrett, P. 1993. “Key Elements in Evaluation” in Department of Finance, *Evaluation: Improving Program Performance and Decision Making.* December. Department of Finance. Canberra. pp19-29.

Battin, T. 1993. “A Break From the Past: The Labor Party and the Political Economy of Keynesian Social Democracy”. *Australian Journal of Political Science.* 28 (2) July. pp221-241.

Baume, P. 1991. “Getting the Policy Right: Recovering the Golden Rules of Evaluation” in J. Uhr (ed.), *Decision Making in Australian Government: Program Evaluation.* Federalism Research Centre, ANU. Canberra. pp35-38.

Beilharz, P. 1994. *Transforming Labor: Labour Tradition and the Labor Decade in Australia.* Cambridge University Press. Melbourne.

Bell, S. & Head, B. 1994. “Australia’s Political Economy: Critical Themes and Issues” in S. Bell & B. Head (eds), *State, Economy and Public Policy in Australia.* Oxford University Press. Melbourne. pp1-24.

Bemelmans-Videc, M. L. 1992. “Institutionalising Evaluation: International Perspectives” in J. Mayne, M. L. Bemelmans-Videc, J. Hudson & R. Connor (eds), *Advancing Public Policy Evaluation: Learning from International Experiences.* North-Holland. Amsterdam. pp7-20.

Bemelmans-Videc, M. L., Eriksen, B. & Goldenberg, E. N. 1994. “Facilitating Organizational Learning: Human Resource Management and Program Evaluation” in F. L. Leeuw, R. C. Rist & R. C. Sonnichsen (eds), *Can Governments Learn? Comparative Perspectives on Evaluation and Organizational Learning.* Transaction Publishers. New Brunswick. pp145-187.

Boston, J. 1989. "The Future Structure and Role of Central Agencies". *Public Sector—Publication of the New Zealand Institute of Public Administration.* 12 (4) December. pp6-10.

Boston, J. 1991. "The Theoretical Underpinnings of Public Sector Restructuring in New Zealand" in J. Boston, J. Martin, J. Pallot & P. Walsh (eds), *Reshaping the State: New Zealand's Bureaucratic Revolution.* Oxford University Press. Auckland. pp1-26.

Boston, J. & Uhr, J. 1996. "Reshaping the Mechanics of Government" in F. G. Castles, R. Gerritsen & J. Vowles (eds), *The Great Experiment: Labour Parties and Public Policy Transformation in Australia and New Zealand.* Allen & Unwin. Sydney. pp48-67.

Bowden, J., Ballard, K. & FitzGerald, R. 1991. "Implementation of Program Evaluation in the Australian Public Service". *Evaluation Journal of Australasia.* 3 (2). pp3-10.

Bulmer, M. 1988. "Social Science Expertise and Executive-Bureaucratic Politics in Britain". *Governance—An International Journal of Policy and Administration.* 1 (1) January. pp26-49.

Campbell, C. SJ. 1983. *Governments Under Stress: Political Executives and Key Bureaucrats in Washington, London and Ottawa.* University of Toronto Press. Toronto.

Campbell, C. SJ. & Halligan, J. 1992. *Political Leadership in an Age of Constraint: Bureaucratic Politics Under Hawke and Keating.* Allen & Unwin. St. Leonards.

Campbell, C. SJ. & Szablowski, G. J. 1979. *The Superbureaucrats: Structure and Behaviour in Central Agencies.* Macmillan. Toronto.

Castles, F. G. 1988. *Australian Public Policy and Economic Vulnerability.* Allen & Unwin. Sydney.

Castles, F. G. 1989. "Social Protection by Other Means: Australia's Strategy of Coping with External Vulnerability" in F. G. Castles (ed), *The Comparative History of Public Policy.* Polity Press. Oxford. pp16-55.

Castles, F. G. 1993. "Changing Course in Economic Policy: The English-Speaking Nations in the 1980s" in F. G. Castles (ed), *Families of Nations: Patterns of Public Policy in Western Democracies.* Dartmouth. Aldershot.

Castles, F. G., Gerritsen, R. & Vowles, J. 1996. "Introduction: Setting the Scene for Economic and Political Change" in F. G. Castles, R. Gerritsen & J. Vowles (eds), *The Great Experiment: Labour Parties and Public Policy Transformation in Australia and New Zealand.* Allen & Unwin. Sydney. pp1-21.

Caulley, D. N. 1993. "What is Problematic About Government Program Evaluation in Australasia Today?". *Evaluation News and Comment: Magazine of the Australasian Evaluation Society.* 2 (1) March. pp2-8.

Chelimsky, E. 1989. "Old Patterns and New Directions in Program Evaluation" in E. Chelimsky (ed), *Program Evaluation: Patterns and Directions.* Second Edition. APSA. Washington DC. pp1-35.

Codd, M. 1990. "Cabinet Operations of the Australian Government" in B. Galligan, J. Nethercote & C. Walsh (eds), *Decision Making in Australian Government: The Cabinet and Budget Processes.* Centre for Research on Federal Financial Relations, ANU. Canberra. pp1-22.

Colebatch, H. K. 1995. "Organizational Meanings of Program Evaluation". *Policy Sciences.* 28 (2). pp149-164.

Corbett, D. 1991. "Contours Since Coombs: The Changing Landscape in Program Evaluation" in J. Uhr (ed.), *Decision Making in Australian Government: Program Evaluation.* Federalism Research Centre, ANU. Canberra. pp3-18.

Craswell, E. & Davis, G. 1993. "Does the Amalgamation of Government Agencies Produce Better Policy Coordination?" in P. Weller, J. Forster & G. Davis (eds), *Reforming the Australian Public Service: Lessons From Recent Experience.* Macmillan. South Melbourne. pp180-207.

Cutt, J. 1977. "Accountability, Efficiency and the Royal Commission on Australian Government Administration". *Australian Journal of Public Administration.* 36 (4) December. pp333-349.

Davidson, R. & White, P. 1988. "Introduction" in R. Davidson & P. White (eds), *Information and Government: Studies in the Dynamics of Policy-Making.* Edinburgh University Press. Edinburgh. pp1-13.

Davis, D. F. 1990. "Do You Want a Performance Audit or a Program Evaluation?". *Public Administration Review.* 50 (1) January/February. pp35-41.

Davis, G. 1995. *A Government of Routines: Executive Coordination in an Australian State.* Macmillan. South Melbourne.

Davis, G., Wanna, J., Warhurst, J. & Weller, P. 1993. *Public Policy in Australia.* Second Edition. Allen & Unwin. St Leonards.

Dawkins, J. S. 1985. "Reforms in the Canberra System of Public Administration". *Australian Journal of Public Administration.* 44 (1) March. pp59-72.

Derlien, H-U. 1990. "Genesis and Structure of Evaluation Efforts in Comparative Perspective" in R. C. Rist (ed), *Program Evaluation and the Management of Government: Patterns and Prospects Across Eight Nations.* Transaction Publishers. New Brunswick. pp147-175.

Dery, D. 1990. *Data and Policy Change: The Fragility of Data in the Policy Context.* Kluwer Academic. Boston.

Dickenson, M. 1984. "Personnel Management and Industrial Relations in the Public Service" in A. Kouzmin, J. R. Nethercote & R. Wettenhall (eds), *Australian Commonwealth Administration. 1983 Essays in Review.* Canberra CAE/RAIPA. Canberra. pp43-59.

Dillon, M. C. 1985. "Developments in Public Sector Audit: How Effective Are Efficiency Audits?". *Australian Journal of Public Administration.* 44 (3) September. pp247-269.

Dobell, R. & Zussman, D. 1981. "An Evaluation System for Government: If Politics is Theatre, Then Evaluation is (Mostly) Art". *Canadian Public Administration.* 24 (3) Fall. pp404-427.

Dollery, B. & Hamburger, P. 1996. "Modelling Bureaucracy: The Case of the Australian Federal Budget Sector 1982-1992". *Public Administration.* 74 Autumn. pp477-507.

Dunleavy, P. 1986. "Explaining the Privatization Boom: Public Choice Versus Radical Approaches". *Public Administration.* 64 Spring. pp13-34.

Dunleavy, P. 1991. *Democracy, Bureaucracy and Public Choice: Economic Explanations in Political Science.* Harvester Wheatsheaf. London.

Dunleavy, P. & Rhodes, R. A. W. 1990. "Core Executive Studies in Britain". *Public Administration.* 68 Spring. pp3-28.

Dunn, D. D. 1995. "Ministerial Staff in Australian Commonwealth Government". *Australian Journal of Public Administration.* 54 (4) December. pp507-519.

Dunn, W. N. & Holzer, B. 1988. "Knowledge in Society: Anatomy of an Emergent Field". *Knowledge in Society.* 1 (1) Spring. pp5-26.

Dunsire, A. 1990. "Holistic Governance". *Public Policy and Administration.* 5 (1) Spring. pp4-19.

Dunsire, A. 1991. "A Cybernetic View of Guidance, Control and Evaluation in the Public Sector" in F-X, Kaufmann (ed), *The Public Sector: Challenge for Coordination and Learning.* Walter de Gruyter. Berlin. pp325-345.

Easton, B. & Gerritsen, R. 1996. "Economic Reform: Parallels and Divergences" in F. G. Castles, R. Gerritsen & J. Vowles (eds), *The Great Experiment: Labour Parties and Public Policy Transformation in Australia and New Zealand.* Allen & Unwin. Sydney. pp22-47.

Edwards, M. 1992. "Evaluating Policy Advice: A Comment". *Australian Journal of Public Administration.* 51 (4) December. pp447-449.

Edwards, M. & Stuart, A. 1994. "The 1994 Employment White Paper: A Preliminary Review of a Policy Process" in *Evaluation: Making Performance Count.* Proceedings of the International Conference of the Australasian Evaluation Society. 22pp.

Edwards, M. 1996. "Lessons From Prime Minister and Cabinet" in J. Uhr & K. Mackay (eds), *Evaluating Policy Advice: Learning From Commonwealth Experience.* Federalism Research Centre, Australian National University and the Commonwealth Department of Finance. Canberra. pp73-82.

Emy, H. V. 1976a. "The Public Service and Political Control—Appendix 1.B" in Royal Commission on Australian Government Administration, *Appendix - Volume One.* AGPS. Canberra. pp15-63.

Emy, H. V. 1976b. *Public Policy: Problems and Paradoxes.* Macmillan. South Melbourne.

Emy, E. V. & Hughes, O. E. 1988. *Australian Politics: Realities in Conflict.* Macmillan. South Melbourne.

Fielding, N. 1993. "Qualitative Interviewing" in N. Gilbert (ed), *Researching Social Life.* Sage Publications. London. pp135-153.

Forrester, G. 1993. "Policy Advising Programs: The Impact of Audit and Evaluation". *Canberra Bulletin of Public Administration.* No. 73. September. pp160-163.

Forsyth, P. 1992. "A perspective on microeconomic reform" in P. Forsyth (ed), *Microeconomic Reform in Australia.* Allen & Unwin. St Leonards. pp3-23.

Fraser, B. W. 1984. "The Treasury—Tendering Economic Advice". *Canberra Bulletin of Public Administration.* XI (4) Summer. pp230-233.

Fry, G. K. 1993. *Reforming the Civil Service: The Fulton Committee on the British Home Civil Service 1966-1968.* Edinburgh University Press. Edinburgh.

Goldring, J. 1981. "The Foundations of the 'New Administrative Law'. *Australian Journal of Public Administration.* 40 (2) June. pp79-102.

Gray, A. & Jenkins, B. 1982. "Policy Analysis in British Government: The Experience of PAR". *Public Administration.* 60 (4). pp429-450.

Gray, A. & Jenkins, B. 1985. *Administrative Politics in British Government.* Harvester Wheatsheaf. Brighton.

Gray, A. & Jenkins, B. 1989. "Policy Evaluation in a Time of Fiscal Stress: Some Reflections from British Experience". *Knowledge in Society.* 2 (4). pp19-30.

Gray, A. & Jenkins, B. 1992. "Implementing Evaluation: Lessons from the UK" in J. Mayne, M. L. Bemelmans-Videc, J. Hudson & R. Connor (eds), *Advancing Public Policy Evaluation: Learning from International Experiences.* North-Holland. Amsterdam. pp59-69.

Gray, A. & Jenkins, B. 1993. "Horses to Water: Budgeting, Auditing and Evaluation in Seven Governments" in A. Gray, B. Jenkins & B. Segsworth (eds), *Budgeting, Auditing and Evaluation: Functions and Integration in Seven Governments.* Transaction Publishers. New Brunswick. pp185-208.

Gray, A., Jenkins, B. & Segsworth, B. 1993. "Perspectives on Budgeting, Auditing and Evaluation: An Introduction" in A. Gray, B. Jenkins & B. Segsworth (eds), *Budgeting, Auditing and Evaluation: Functions and Integration in Seven Governments.* Transaction Publishers. New Brunswick. pp1-18.

Gregory, P. 1996. "Policy Management Reviews: Summary of Findings" in J. Uhr & K. Mackay (eds), *Evaluating Policy Advice: Learning From Commonwealth Experience.*

Federalism Research Centre, Australian National University and the Commonwealth Department of Finance. Canberra. pp140-160.

Gregory, R. 1995. “Accountability, Responsibility and Corruption: Managing the ‘Public Production Process’” in J. Boston (ed), *The State Under Contract.* Bridget Williams Books. Wellington. pp56-77.

Gross, P. 1979a. “Recent Evaluation Activity in Australia” in SSCSW, *Through a Glass, Darkly: Evaluation in Australian Health and Welfare Services* Volume 2: Papers commissioned by the Committee. AGPS. Canberra. pp41-59.

Gross, P. 1979b. “Parliament, the Executive and External Review: Program Evaluation and the Role of Government and Non-Government Bodies” in *New Developments in Public Sector Management*. Conference Papers and Report. March. Centre for Continuing Education, ANU. Canberra. pp79-104.

Gruen, F. & Grattan, M. 1993. *Managing Government: Labor’s Achievements and Failures.* Longman Cheshire. Melbourne.

Guba, E. G. 1990. “The Alternative Paradigm Dialog” in E. G. Guba (ed), *The Paradigm Dialog.* Sage Publications. Newbury Park. pp17-27.

Halligan, J. 1987. “Reorganising Australian Government Departments”. *Canberra Bulletin of Public Administration.* 52 October. pp40-47.

Halligan, J. 1988. “The Australian Public Service Reform Program” in R. Wettenhall & J. R. Nethercote (eds), *Hawke’s Second Government: Australian Commonwealth Administration 1984-1987.* Canberra CAE / RIPAA. Canberra. pp27-84.

Halligan, J. 1994. “The Process of Reform: Balancing Principle and Pragmatism” in J. Stewart (ed), *From Hawke to Keating: Australian Commonwealth Administration 1990-1993.* University of Canberra / RIPAA. Canberra. pp1-14.

Halligan, J., Beckett, I. & Earnshaw, P. 1992. “The Australian Public Service Reform Program” in J. Halligan & R. Wettenhall (eds), *Hawke’s Third Government: Australian Commonwealth Administration 1987-1990.* University of Canberra / RIPAA (ACT Division). Canberra. pp7-56.

Halligan, J. & Power, J. 1992. *Political Management in the 1990s.* Oxford University Press. Melbourne.

Halligan, J. & Wettenhall, R. 1990. “Major Changes in the Structure of Government Institutions” in J. Power (ed), *Public Administration in Australia: A Watershed.* Hale & Iremonger. Sydney. pp17-40.

Hamburger, P. 1992. “Parliament and Program Evaluation” in C. O'Faircheallaigh and B. Ryan (eds), *Program Evaluation and Performance Monitoring: An Australian Perspective.* Macmillan. South Melbourne. pp77-90.

Hawke, B. 1987. “New Machinery of Government: The Prime Minister’s Media Statement 14 July 1987”. *Canberra Bulletin of Public Administration.* 52 October. pp12-14.

Hawke, B. 1994. *The Hawke Memoirs.* William Heinemann Australia. Melbourne.

Hawker, G. 1981. *Who's Master, Who's Servant? Reforming Bureaucracy*. George Allen and Unwin. Sydney.

Hawker, G., Smith, R.F.I. & Weller, P. 1979. *Politics and Policy in Australia.* University of Queensland Press. St Lucia.

Heclo, H. & Wildavsky, A. 1974. *The Private Government of Public Money: Community and Policy Inside British Politics.* Macmillan. London.

Hellstern, G-M. 1991. “Generating Knowledge and Refining Experience: The Task of Evaluation” in F-X. Kaufmann (ed), *The Public Sector: Challenge for Coordination and Learning.* Walter de Gruyter. New York. pp271-307.

Henkel, M. 1991. *Evaluation and Change.* Jessica Kingsley Publishers. London.

Hicks, M. 1981. “‘Evaluating Evaluation in Today’s Government’: Summary of Discussions”. *Canadian Public Administration.* 24 (3) Fall. pp350-358.

Higgins, C. & Borthwick, D. 1988. “The Treasury”. *Economic Papers.* 7 (4) December. pp60-74.

Higgins, C. & Borthwick, D. 1990. “The Role of Central Agencies” in B. Galligan, J. R. Nethercote & C. Walsh (eds), *Decision Making in Australian Government: The Cabinet and Budget Processes.* Centre for Research on Federal Financial Relations, ANU. Canberra. pp43-61.

Holmes, M. & Shand, D. 1995. “Management Reform: Some Practitioner Perspectives on the Past Ten Years”. *Governance—An International Journal of Policy and Administration.* 8 (4). pp551-578.

Hood, C. 1990. “De-Sir Humphreyfying the Westminster Model of Bureaucracy: A New Style of Governance”. *Governance—An International Journal of Policy and Administration.* 3 (2) April. pp205-214.

Hood, C. 1991. “Concepts of Control over Public Bureaucracies: ‘Comptrol’ and ‘Interpolable Balance’” in F-X. Kaufmann (ed), *The Public Sector: Challenge for Coordination and Learning.* Walter de Gruyter. Berlin. pp347-366.

Hood, C. 1994. *Explaining Economic Policy Reversals.* Open University Press. Buckingham.

Hood, C. 1995. “The ‘New Public Management’ in the 1980s: Variations on a Theme”. *Accounting, Organizations and Society.* 20 (2/3). pp93-109.

Howard, J. 1986. “Public Expenditure Management: The Rhetoric of Planning and the Reality of Control” in J. R. Nethercote, A. Kouzmin & R. Wettenhall (eds), *Australian Commonwealth Administration: 1984 Essays in Review*. Canberra CAE/RAIPA. Canberra. pp40-73.

Hughes, O. 1994. *Public Management and Administration: An Introduction.* St Martin’s Press. New York.

Jinks, B. 1982. “The ‘New Administrative Law’: Some Assumptions and Questions”. *Australian Journal of Public Administration.* 41 (3) September. pp209-218.

Johnston, J. 1992. “Program Evaluation and Performance Monitoring in New South Wales” in C. O'Faircheallaigh and B. Ryan (eds), *Program Evaluation and Performance Monitoring: An Australian Perspective.* Macmillan. South Melbourne. pp35-46.

Jordon, J. & Sutherland, S. L. 1979. “Assessing the Results of Public Expenditure: Program Evaluation in the Canadian Federal Government”. *Canadian Public Administration.* 22 (4). pp581-609.

Kaufman, H. 1981. *The Administrative Behaviour of Federal Bureau Chiefs.* The Brookings Institution. Washington DC.

Kaufmann, F-X. 1991a. “Introduction: Issues and Context” in F-X, Kaufmann (ed), *The Public Sector: Challenge for Coordination and Learning.* Walter de Gruyter. Berlin. pp3-28.

Kaufmann, F-X. 1991b. “The Relationship Between Guidance, Control and Evaluation” in F-X. Kaufmann (ed), *The Public Sector: Challenge for Coordination and Learning.* Walter de Gruyter. Berlin. pp213-234.

Keating, M. 1989. “Quo Vadis? Challenges of Public Administration”. *Australian Journal of Public Administration.* 48 (2). pp123-131.

Keating, M. 1990. “Managing for Results in the Public Interest”. *Australian Journal of Public Administration.* 49 (4). pp387-398.

Keating, M. 1990. “The Process of Commonwealth Budgetary Control” in J. Forster & J. Wanna (eds), *Budgetary Management and Control: The Public Sector in Australasia.* Macmillan. South Melbourne. pp1-16.

Keating, M. 1993a. “Mega-departments: The Theory, Objectives and Outcomes of the 1987 Reforms” in P. Weller, J. Forster & G. Davis (eds), *Reforming the Australian Public Service: Lessons From Recent Experience.* Macmillan. South Melbourne. pp1-13.

Keating, M. 1993b. “The Role of the Department of the Prime Minister and Cabinet in the Policy Making Process: Adding Value”. Paper presented at the AIC Conference on *Developing, Managing, Implementing and Evaluating Government Policy: The Quest for Quality*. Sheraton Wentworth Hotel. Sydney. 17-18 May.

Keating, M. 1994. “The Role of Government Economists”. *Canberra Bulletin of Public Administration*. 77 December. pp1-7.

Keating, M. 1995. “The Role of the Public Service in Policy Development”. Paper presented at the AIC *Government Policy Conference.* Intercontinental Hotel. Sydney. 13 June.

Keating, M. & Holmes, M. 1990. "Australia's Budgetary and Financial Management Reforms". *Governance—An International Journal of Policy and Administration.* 3 (2) April. pp168-185.

Keating, P. J. 1992. *One Nation.* Speech by the Prime Minister The Honourable P. J. Keating, MP. 26 February 1992. AGPS. Canberra.

Kelly, P. 1994. *The End of Certainty: Power, Politics and Business in Australia.* Revised Edition. Allen & Unwin. St Leonards.

Kennedy, P. 1988. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000.* Unwin Hyman. London.

Lester, J. P. & Wilds, L. J. 1990. "The Utilization of Public Policy Analysis: A Conceptual Framework". *Evaluation and Program Planning.* 13. pp313-319.

Levine, R. A. 1981. "Program Evaluation and Policy Analysis in Western Nations: An Overview" in R. A. Levine, M. A. Solomon, G. M. Hellstern & H. Wollmann (eds), *Evaluation Research and Practice: Comparative and International Perspectives.* Sage Publications. London. pp27-60.

Lindblom, C. E. 1965. *The Intelligence of Democracy.* The Free Press. New York.

Lindquist, E. A. 1988. "What Do Decision Models Tell Us About Information Use?". *Knowledge in Society.* 1 (2) Summer. pp86-111.

Machiavelli, N. 1995 [1961]. *The Prince.* Translated with an introduction by George Bull. Penguin Books. London.

Mackay, K. 1993. "Federal Initiatives: The Evolution of Evaluation in the Commonwealth Government" in Department of Finance, *Evaluation: Improving Program Performance and Decision Making.* December. Department of Finance. Canberra. pp9-18.

Mackay, K. 1994a. "The Australian Government's Evaluation Strategy: A Perspective From the Center". *Canadian Journal of Program Evaluation.* 9 (2). pp15-30.

Mackay, K. 1994b. "The Commonwealth Government's Evaluation Strategy" in SSCFPA (Senate Standing Committee on Finance and Public Administration), *Public Service Reform. Volume 2: Conference Proceedings.* August. Department of the Senate. Canberra. pp225-240.

Mackay, K. 1996. "The Institutional Framework for Evaluation in the Australian Government". Paper delivered to a World Bank seminar *Australia's Program and Policy Evaluation Framework: Institutionalising the Process.* Washington DC. 30 April.

Manning, H. 1992. "The ALP and the Union Movement: 'Catch-All' Party or Maintaining Tradition". *Australian Journal of Political Science.* 27 (1) March. pp12-30.

Martin, J. 1995. "Contracting and Accountability" in J. Boston (ed), *The State Under Contract.* Bridget Williams Books. Wellington. pp36-55.

Mayne, J. & Mayne, R. S. 1983. "Will Program Evaluation Be Used In Formulating Policy?" in M. M. Atkinson & M. A. Chandler (eds), *The Politics of Canadian Public Policy*. University of Toronto Press. Toronto. pp267-283.

McQueen, C. 1992. "Program Evaluation in the Canadian Federal Government" in J. Hudson, J. Mayne & R. Thomlison (eds), *Action-Oriented Evaluation in Organizations: Canadian Practices.* Wall & Emerson. Toronto. pp28-47.

Meltsner, A. J. 1990. *Rules for Rulers: The Politics of Advice.* Temple University Press. Philadelphia.

Mills, S. 1992. "The Great Game". *Australian Financial Review Magazine.* June. pp22-32.

Minnery, J. R. 1988. "Modelling Coordination". *Australian Journal of Public Administration.* 47 (3) September. pp253-262.

Mosher, F. C. 1979. *The GAO: The Quest for Accountability in American Government.* Westview Press. Boulder, Colorado.

Nethercote, J. R. 1977. "Efficient Allocation of Resources Within the Public Service" in C. Hazlehurst & J. R. Nethercote (eds), *Reforming Australian Government: The Coombs Report and Beyond.* RIPAA (ACT) & ANU Press. Canberra. pp95-118.

Nethercote, J. R. 1984. "Public Service Reform: Its Course and Nature" in A. Kouzmin, J. R. Nethercote & R. Wettenhall (eds), *Australian Commonwealth Administration. 1983 Essays in Review.* Canberra CAE/RIPAA Canberra. pp16-42.

Nethercote, J. R. 1988. "Changing Climate of Public Service" in R. Wettenhall & J. R. Nethercote (eds), *Hawke's Second Government: Australian Commonwealth Administration 1984-1987.* CCAE/RAIPA (ACT Division). Canberra. pp1-26.

Painter, M. 1981. "Central Agencies and the Coordination Principle". *Australian Journal of Public Administration.* 40 (4) December. pp265-280.

Painter, M. 1987. *Steering the Modern State: Changes in Central Coordination in Three Australian State Governments.* Sydney University Press. Sydney.

Painter, M. & Carey, B. 1979. *Politics Between Departments: The Fragmentation of Executive Control in Australian Government.* University of Queensland Press. St Lucia.

Parker, L. D. 1990. "Towards Value-for Money Audit Policy" in J. Guthrie, L. Parker and D. Shand (eds), *The Public Sector: Contemporary Readings in Accounting and Auditing.* Harcourt Brace Jovanovich. Sydney. pp292-306.

Patton, M. Q. 1990. *Qualitative Evaluation and Research Methods.* Second Edition. Sage Publications. Newbury Park.

Peres, L. 1974. "The Politics of Industrial Policy" in AIPS (Australian Institute of Political Science), *Industrial Policy 1975-2000: Preparing for Change.* Proceedings of the 40th Summer School. ANZ Book Company. Sydney. pp148-165.

Peres, L. 1978. "So where's the coordinating capacity coming from?" in R. F. I Smith & P. Weller (eds), *Public Service Inquiries in Australia.* University of Queensland Press. St Lucia. pp249-254.

Peters, B Guy. & Barker, A. 1993. "Introduction: Governments, Information, Advice and Policy-making" in B. Guy Peters & A. Barker (eds), *Advising West European Governments: Inquiries, Expertise and Public Policy*. Edinburgh University Press. Edinburgh. pp1-19.

Pollitt, C. 1993. "Occasional Excursions: A Brief History of Policy Evaluation in the UK". *Parliamentary Affairs.* 46 (3) July. pp353-362.

Pollitt, C. 1995. "Justification by Works or by Faith? Evaluating the New Public Management". *Evaluation: The International Journal of Theory, Research and Practice.* 1 (2). pp133-154.

Prasser, S. 1988. "Public Inquiries Under the Second Hawke Government" in R. Wettenhall & J. R. Nethercote (eds), *Hawke's Second Government: Australian Commonwealth Administration 1984-1987.* CCAE/RAIPA (ACT Division). Canberra. pp117-142.

Pressman, J. L. & Wildavsky, A. 1979. *Implementation.* Second Edition. University of California Press. Berkeley.

Pusey, M. 1991. *Economic Rationalism in Canberra: A Nation Building State Changes Its Mind.* Cambridge University Press. Melbourne.

Reid, G. S. 1976. "Responsibility and Accountability and the Coombs Inquiry". *Australian Journal of Public Administration.* 45 (4) December. pp320-327.

Reid, G. S. & Forrest, M. 1989. *Australia's Commonwealth Parliament 1901-1988: Ten Perspectives.* Melbourne University Press. Carlton, Victoria.

Rist, R. C. 1989. "Management Accountability: The Signals sent by Auditing and Evaluation". *Journal of Public Policy*. 9 (3). pp355-369.

Rist, R. C. 1990. "Managing of Evaluations or Managing by Evaluations: Choices and Consequences" in R. C. Rist (ed), *Program Evaluation and the Management of Government: Patterns and Prospects Across Eight Nations.* Transaction Publishers. New Brunswick. pp3-17.

Rogers, V. 1993. "Australian Public Service Reform: Managing a 1990s-style Inquiry". *Australian Journal of Public Administration.* 52 (4) December. pp371-375.

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. pp63-76.

Savoie, D. J. 1990. *The Politics of Public Spending in Canada.* University of Toronto Press. Toronto.

Savoie, D. J. 1994. *Thatcher, Reagan, Mulroney: In Search of a New Bureaucracy.* University of Toronto Press. Toronto.

Savoie, D. J. 1995. "Globalisation, Nation States, and the Civil Service" in B. Guy Peters & D. J. Savoie (eds), *Governance in a Changing Environment*. McGill-Queen's University Press. Montreal & Kingston. pp82-110.

Schick, A. 1986a. "Macro-Budgetary Adaptations to Fiscal Stress in Industrialized Democracies". *Public Administration Review*. 46 (2) March / April. pp124-134.

Schick, A. 1986b. "Incremental Budgeting in a Decremental Age" in F. S. Lane (ed), *Current Issues in Public Administration*. Third Edition. St Martin's Press. New York. pp276-299.

Schick, A. 1988. "Micro-Budgetary Adaptations to Fiscal Stress in Industrialized Democracies". *Public Administration Review*. 48 (1) January / February. pp523-533.

Schick, A. 1990. "Budgeting for Results: Recent Developments in Five Industrialized Countries". *Public Administration Review*. 50 (1) January / February. pp26-34.

Schwartz, H. M. 1994a. "Public Choice Theory and Public Choices: Bureaucrats and State Reorganization in Australia, Denmark, New Zealand, and Sweden in the 1980s". *Administration and Society*. 26 (1) May. pp48-77.

Schwartz, H. M. 1994b. "Small States in Big Trouble: State Reorganization in Australia, Denmark, New Zealand, and Sweden in the 1980s". *World Politics*. 46 July. pp527-55.

Sedgwick, S. 1992. "Encouraging Evaluation in the Australian Budget Sector" in J. Mayne, M. L. Bemelmans-Videc, J. Hudson & R. Connor (eds), *Advancing Public Policy Evaluation: Learning From International Experiences*. North Holland. Amsterdam. pp37-48.

Sedgwick, S. 1993. "The Role of Evaluation in the Budget" in Department of Finance, *Evaluation: Improving Program Performance and Decision Making*. December. Department of Finance. Canberra. pp1-8.

Sedgwick, S. 1994. "Finance—Improving the Way That it Does Business". Presentation to the Department of Finance Briefings for Senior Managers, *Trends in Public Sector Financial Management*. Lakeside Hotel, Canberra. 14 July. 14pp.

Segsworth, B. 1990. "Policy and Program Evaluation in the Government of Canada" in R. C. Rist (ed), *Program Evaluation and the Management of Government: Patterns and Prospects Across Eight Nations*. Transaction Publishers. New Brunswick. pp21-36.

Segsworth, B. 1993. "Out of Sequence and Out of Sync: Budgeting, Auditing and Evaluation in Canada" in A. Gray, B. Jenkins & B. Segsworth (eds), *Budgeting, Auditing and Evaluation: Functions and Integration in Seven Governments*. Transaction Publishers. New Brunswick. pp87-114.

Self, P. 1978. "The Coombs Commission: An Overview" in R. F. I. Smith & P. Weller (eds), *Public Service Inquiries in Australia*. University of Queensland Press. St Lucia. pp310-333.

Self, P. 1980. “Resource and Policy Coordination under Pressure” in R. Rose (ed), *Challenge to Governance: Studies in Overloaded Polities.* Sage Publications. London. pp29-51.

Self, P. 1984. “Administrative Overload and the Role of Government”. *Australian Journal of Public Administration.* 43 (1) March. pp12-19.

Self, P. 1993. *Government by the Market? The Politics of Public Choice.* Macmillan. London.

Shadish, W. R., Cook, T. D. & Leviton, L. C. 1991. *Foundations of Program Evaluation: Theories of Practice.* Sage Publications. Newbury Park.

Sharman, C. & Stuart, J. 1982. “Premier’s Departments: Patterns of Growth and Change”. *Politics.* 17 (1) May. pp46-58.

Singleton, G. 1990. *The Accord and the Australian Labour Movement.* Melbourne University Press. Melbourne.

Smith, M. J., Marsh, D. & Richards, D. 1993. “Central Government Departments and the Policy Process”. *Public Administration.* 71 Winter. pp567-594.

Smith, R. F. I. 1978. “Central Coordination and Control” in R. F. I Smith & P. Weller (eds.), *Public Service Inquiries in Australia.* University of Queensland Press. St Lucia. pp177-192.

Spann, R. 1977. “The Coombs Doctrine” in C. Hazelhurst & J. Nethercote (eds), *Reforming Australian Government: The Coombs Report and Beyond.* RIPA/ANU Press. Canberra. pp78-86.

Spann, R. N. 1979. *Government Administration in Australia.* Allen & Unwin. Sydney.

Stewart, J. 1992. “Corporatism, Pluralism and Political Learning: A Systems Approach”. *Journal of Public Policy.* 12 (3). pp243-255.

Stone, B. 1994. “Success in Public Inquiries: An Analysis and a Case Study” in P. Weller (ed.), *Royal Commissions and the Making of Public Policy.* Macmillan. South Melbourne. pp244-258.

Strickland, A. J. 1982. “South Australia’s Program-Performance Budgeting Experience” in J. R. Nethercote (ed), *Parliament and Bureaucracy.* Hale & Iremonger. Sydney. pp113-124.

Stutchbury, M. 1990. “Macroeconomic Policy” in C. Jennett & R. G. Stewart (eds), *Hawke and Australian Public Policy: Consensus and Restructuring.* Macmillan. South Melbourne. pp54-78.

Swank, D. H. 1992. “Electoral and Partisan Influences on Australian Fiscal Policy From Menzies to Hawke”. *Australian Journal of Political Science.* 27 (3) November. pp414-433.

Tarschys, D. 1983. “Fluctuations in the Political Demand for Policy Analysis” in S. E. Spiro & E. Yuchtman-Yaar (eds), *Evaluating the Welfare State: Social and Political Perspectives.* Academic Press. New York. pp375-385.

Thain, C. & Wright, M. 1992a. “Planning and Controlling Public Expenditure in the UK, Part I: The Treasury’s Public Expenditure Survey”. *Public Administration.* 70 Spring. pp3-24.

Thain, C. & Wright, M. 1992b. “Planning and Controlling Public Expenditure in the UK, Part II: The Effects and Effectiveness of the Survey”. *Public Administration.* 70 Summer. pp193-224.

Thompson, E. 1979. “The Public Service” in A. Patience & B. Head (eds), *From Whitlam to Fraser: Reform and Reaction in Australian Politics.* Oxford University Press. Melbourne. pp70-87.

Thompson, E. 1989. “The Public Service” in B. W. Head & A. Patience (eds), *From Fraser to Hawke.* Longman Cheshire. Melbourne. pp213-239.

Uhr, J. 1982. “Parliament and Public Administration” in J. R. Nethercote (ed), *Parliament and Bureaucracy.* Hale & Iremonger. Sydney. pp26-66.

Uhr, J. 1990a. “Public Expenditure and Parliamentary Accountability: The Debatable Role of Senate Estimates Committees” in *Senate Estimates Scrutiny of Government Finance and Expenditure. What's it For, Does it Work and at What Cost? Papers on Parliament.* No. 6. March. Department of the Senate. Canberra. pp79-107.

Uhr, J. 1990b. “The Work of Evaluation and the Evaluation of Work”. *Canberra Bulletin of Public Administration.* No. 63. December. pp89-92.

Uhr, J. 1996. “Lessons From an External Reviewer” in J. Uhr & K. Mackay (eds), *Evaluating Policy Advice: Learning From Commonwealth Experience.* Federalism Research Centre, Australian National University and the Commonwealth Department of Finance. Canberra. pp101-110.

Uhr, J. 1996 (forthcoming). “Testing the Policy Capacities of Budgetary Agencies: Lessons From Finance”. *Australian Journal of Public Administration.*

Uhr, J. & Di Francesco, M. 1994. “Evaluation and Reporting Reforms” in *Public Service Reform: Volume 1—Conference Papers.* Report from the Standing Committee on Finance and Public Administration. Department of the Senate. Canberra. pp241-267.

Uhr, J. & Mackay, K. 1992. “Trends in Program Evaluation: Guest Editors’ Introduction”. *Australian Journal of Public Administration.* 51 (4) December. pp433-435.

Uhr, J. & Mackay, K. (eds). 1996. *Evaluating Policy Advice: Learning From Commonwealth Experience.* Federalism Research Centre, Australian National University and the Commonwealth Department of Finance. Canberra.

Van Loon, R. 1983. “The Policy and Expenditure Management System in the Federal Government: The First Three Years”. *Canadian Public Administration.* 26 (2) Summer. pp255-285.

Wagner, P. & Wollmann, H. 1986. “Fluctuations in the Development of Evaluation Research: Do ‘Regime Shifts’ Matter?”. *International Social Science Journal.* 108. pp205-218.

Waller, M. 1992. "Evaluating Policy Advice". *Australian Journal of Public Administration*. 51 (4) December. pp440-446.

Waller, M. 1996. "The Changing Environment of Policy Advising" in J. Uhr & K. Mackay (eds), *Evaluating Policy Advice: Learning From Commonwealth Experience*. Federalism Research Centre, ANU and the Commonwealth Department of Finance. Canberra. pp69-72.

Walsh, C. 1991a. "The Politics and Economics of Budgeting Reform" in J. Uhr (ed), *Decision Making in Australian Government: Program Evaluation*. Federalism Research Centre, ANU. Canberra. pp19-34.

Walsh, C. 1991b. "The National Economy and Management Strategies" in B. Galligan & G. Singleton (eds), *Business and Government Under Labor*. Longman Cheshire. Melbourne. pp35-46.

Walsh, P. 1995. *Confessions of a Failed Finance Minister*. Random House Australia. Milsons Point.

Walter, J. 1986. *The Ministers' Minders: Personal Advisers in National Government*. Oxford University Press. Melbourne.

Walter, J. 1992. "Prime Ministers and their Staff" in P. Weller (ed), *Menzies to Keating: The Development of the Australian Prime Ministership*. Melbourne University Press. Carlton. pp28-63.

Wanna, J., Kelly, J. & Forster, J. 1996. "The Rise and Rise of the Department of Finance". Paper presented to the Australasian Political Studies Conference. University of Western Australia. Perth. October 2-4.

Warhurst, J. 1988. "Reforming Central Government Administration in Australia" in C. Campbell & B. Guy Peters (eds), *Organizing Governance, Governing Organizations*. University of Pittsburgh Press. Pittsburgh. pp327-343.

Warhurst, J. 1992. "The Department of Prime Minister and Cabinet in Australia, 1972-1990". *The Round Table*. 324. pp501-513.

Weaver, R. Kent & Rockman, B. A. 1993. "Assessing the Effects of Institutions" in R. Kent Weaver & B. A Rockman (eds), *Do Institutions Matter? Government Capabilities in the United States and Abroad*. The Brookings Institution. Washington, DC. pp1-41.

Weller, P. 1976. "The Public Expenditure Committee: A Parliamentary Experiment", *Newsletter of the Royal Institute of Public Administration (ACT Group)*. III (3) August. pp36-9.

Weller, P. 1977a. "Splitting the Treasury: Old Habits in New Structures?". *The Australian Quarterly*. 49 (1) March. pp29-39.

Weller, P. 1977b. *The Treasury and the Politics of Advice*. Public Policy Paper No. 9. Department of Political Science. The University of Tasmania. July. pp15.

Weller, P. 1983. “Transition: Taking Power in 1983”. *Australian Journal of Public Administration*. 42 (3) September. pp303-319.

Weller, P. 1989. *Malcolm Fraser PM: A Study in Prime Ministerial Power in Australia.* Penguin. Ringwood.

Weller, P. 1990. “The Cabinet” in C. Jennett & R. G. Stewart (eds), *Hawke and Australian Public Policy: Consensus and Restructuring.* Macmillan. South Melbourne. pp16-26.

Weller, P. 1992. “Prime Ministers and Cabinet” in P. Weller (ed), *Menzies to Keating: The Development of the Australian Prime Ministership.* Melbourne University Press. Carlton. pp5-27.

Weller, P. 1996. “Commonwealth-State Reform Processes: A Policy Management Review”. *Australian Journal of Public Administration.* 55 (1) March. pp95-110.

Weller, P & Cutt, J. 1976. *Treasury Control in Australia: A Study in Bureaucratic Politics.* Ian Novak. Sydney.

Weller, P. & Grattan, M. 1981. *Can Ministers Cope? Australian Federal Ministers at Work.* Hutchinson. Richmond, Victoria.

Weller, P. & Lewis, C. 1989. “Corporate Management: Background and Dilemmas” in G. Davis, P. Weller & C. Lewis (eds), *Corporate Management in Australian Government: Reconciling Accountability and Efficiency.* Macmillan. South Melbourne. pp1-16.

Weiss, C. H. 1983. “Ideology, Interests, and Information: The Basis of Policy Decisions” in D. Callahan & B. Jennings (eds), *Ethics, The Social Sciences and Policy Analysis.* Plenum Press. New York. pp213-245.

Widavsky, A. 1979. *The Art and Craft of Policy Analysis.* Macmillan. London.

Wildavsky, A. 1984. *The Politics of the Budgetary Process.* Fourth Edition. Little, Brown & Company. Boston.

Wilenski, P. 1986. *Public Power and Public Administration.* Hale and Iremonger. Sydney.

Wright, M. 1981. “Big Government in Hard Times: The Restraint of Public Expenditure” in M. Wright & C. Hood (eds), *Big Government in Hard Times.* Martin Robertson. Oxford. pp3-31.

Zifcak, S. 1994. *New Managerialism: Administrative Reform in Whitehall and Canberra.* Open University Press. Buckingham.

Appendix I — Interviews

Information is presented in the following format: name, position, date of interview, place of interview.

01 *Allen, Maureen.* Director. Evaluation Directorate. Department of the Treasury. 10 May 1995. Canberra.

02 *Barrados, Maria.* Dr. Deputy Auditor General. Office of the Auditor General of Canada. 15 September 1995. Ottawa, Canada.

03 *Barrett, Pat.* Deputy Secretary. Department of Finance. 1 May 1995. Canberra.

04 *Baume, Peter.* Professor, AO. Chairman of the Senate Standing Committee on Social Welfare 1976-1980. Senator for New South Wales, 1974-91. 13 July 1995. Canberra.

05 *Burmetser, Bill.* Assistant Secretary. Employment, Education and Training Branch. Labour and International Division. Department of Finance. 31 May 1995. Canberra.

06 *Corrigan, Margaret.* Industry, Technology and Commerce Branch. Defence and Industry Division. Department of Finance. 5 June 1995. 6 June 1995. Canberra.

07 *Duckett, Stephen.* Dr. Secretary. Department of Human Services and Health. 14 August 1995. Woden, ACT.

08 *Edwards, Meredith.* Dr. Deputy Secretary. Department of the Prime Minister and Cabinet. 30 June 1995. Canberra.

09 *Forrester, Geoff.* Deputy Secretary. Department of Foreign Affairs and Trade. 8 August 1995. Canberra.

10 *Hammond, Cheryl.* Health and General Branch. Social Security Division. Department of Finance. 13 June 1995. Canberra.

11 *Hehir, Grant.* Assistant Secretary. Public Service Employment Branch. Labour and International Division. Department of Finance. 6 June 1995. Canberra.

12 *Keating, Michael.* Dr, AO. Secretary. Department of the Prime Minister and Cabinet. 12 July 1995. Canberra.

13 *Leblanc, Celine.* Director. Audit and Evaluation. Canadian Radio and Telecommunications Commission. 18 September 1995. Hull, Quebec, Canada.

14 *Lennon, Brett.* Assistant Secretary. Health and General Branch. Social Security Division. Department of Finance. 31 May 1995. Canberra.

15 *Mackay, Keith.* Assistant Secretary. Evaluation and Staffing Analysis Branch. Department of Finance. 24 May 1995; 21 June 1995. Canberra.

16 *MacRae, Ian.* Health and General Branch. Social Security Division. Department of Finance. 16 June 1995. 30 June 1995. Canberra.

17 *Mayne, John.* Dr. Office of the Auditor General of Canada. 15 September 1995. Ottawa, Canada.

18 *McMahon, Pat.* Health and General Branch. Social Security Division. Department of Finance. 7 June 1995. Canberra.

19 *Merritt-Balkos, Leanne.* Evaluation and Staffing Analysis Branch. Department of Finance. 15 June 1995. Canberra.

20 *Nicholson, John.* Group Director. Evaluation Consultancy Group. Evaluation and Staffing Analysis Branch. Department of Finance. 24 May 1995. Canberra.

21 *Preston, Alan.* Dr. Deputy Secretary. Fiscal Division. Department of the Treasury. 20 June 1995. Canberra.

22 *Reddan, Nick.* Evaluation and Staffing Analysis Branch. Department of Finance. 30 May 1995. Canberra.

23 *Robertson, Irene.* Evaluation and Staffing Analysis Branch. Department of Finance. 24 May 1995. Canberra.

24 *Russell, Derek.* Outcomes Reporting Group. Resource Management Improvement Branch. Department of Finance. 28 June 1995. Canberra.

25 *Sedgwick, Stephen.* Secretary. Department of Finance. 1 August 1995. Canberra.

26 *Smith, Gary.* Management Review Branch. Department of Finance. 21 June 1995. Canberra.

27 *Smith, Greg.* Acting Deputy Secretary. Financial and Structural Division. Department of the Treasury. 7 June 1995. Canberra.

28 *Ulrich, Martin.* Director of Operations. Government Review and Quality Service Sector. Treasury Board Secretariat. 18 September 1995. Ottawa, Canada.

29 *Wileman, Tom.* Senior Researcher. Office of the Auditor General of Canada. 20 September 1995. Ottawa, Canada.

30 *Winberg, Alan.* Government Review and Quality Services Sector. Treasury Board Secretariat of Canada. 22 September 1995. Ottawa, Canada.

Appendix II

Program Structure For The Commonwealth Department Of Finance

Please Note: to see structure, please see print version.

Appendix III

Questionaire For *Use Of Evaluation In The Budget* Series

Please note: to see questionaire, please see print version.