

**PERFORMANCE REPORTING BY NEW ZEALAND CENTRAL  
GOVERNMENT AGENCIES**

BY  
RAHMADI MURWANTO

A thesis  
submitted to the Victoria University of Wellington  
in fulfilment of the requirements for the degree of  
Doctor of Philosophy

Victoria University of Wellington  
2011

## **ABSTRACT**

This study attempts to assess longitudinally the increase in the quality of performance information resulting from the Managing for Outcome (MfO) initiative, embarked on by the New Zealand Government in December 2001, and to identify the internal and external factors influencing the increase in quality. To measure the quality of performance information, a disclosure index was developed. The index assesses the comprehensiveness of information in compliance with available guidance, and reflects the approaches used by Marston and Shriver (1991), and Guthrie et. al. (2004). The disclosure analysis was applied to publicly available planning documents - the Statement of Intent and Annual Report of 27 New Zealand Government departments over the period 2003-2007.

Agency theory, focusing on the role of information in the accountability relationship between principals and their agents, and public choice theory, focusing on the mechanisms to mitigate public choice problems, are used to explain the improvement in the quality of performance information and the external and internal factors influencing the improvement in quality. The roles performed and the activities initiated and implemented by ministers and other government agencies in the MfO initiative are identified and analysed. The data for the study was obtained from the reports of selected New Zealand central government departments and from semi-structured interviews.

The findings support the Auditor General's assertion of disappointing quality in performance information. Weak incentives for reporting outcomes, the lack of authoritative reporting standards, and constraints on measuring performance have been the key factors in explaining the lack of meaningful progress in New Zealand performance reporting practice implemented under the MfO initiative. The initiatives do not include proper accountability arrangements, where the ministers responsible for

outcomes also report; instead the current arrangement is that chief executives report but are not themselves accountable.

## **Acknowledgment**

Throughout the long and often exasperating path to complete my PhD there have been many people that have greatly supported me to make completion possible. Academics, officials, friends and family have all contributed to this challenging project.

The academics at the School of Accounting and Commercial Law and the School of Government of Victoria University of Wellington's School have given me the benefit of their notable academic experience. I owe my deepest gratitude to my supervisors who have gone well beyond what is expected of doctoral supervisors. Prof. Tony van Zijl always challenged my ideas with his critical inquiry yet at times came out with humorous comments. Dr. Bhagwan Khanna, with his constant support, had been more like a parent, not only for me but also for my family. Both of them have contributed to my intellectual development throughout my PhD journey.

My very special appreciation is also extended to Prof. Kevin Simpkins, Miss RoseAnne MacLeod, Mr Derek Gill and Prof. David Macdonald for their support in getting access to various people during my research, helping me shaping my research instruments, and making sure that my research is in the proper context of the New Zealand public sector.

I am grateful for the support given by the School of Accounting and Commercial Law of Victoria University of Wellington. My special thank is for the current and former Head of School, Prof. Ian Eggleton and Prof. Keitha Dunstan, who have supported my research in many ways.

I would also like to acknowledge my gratitude for the scholarship from NZAID that I have been fortunate to receive and without which much of my research would not have been possible. I would like to thank all NZAID Contract Managers, Miss Bridget Nankivell, Miss Inge de Leeuw, and Miss Julia Harrison, who have ensured that all the support required for me and my family was always available.

I owe a debt to the government officials who gave generously of their time to be interviewed and thus provided important empirical data reported and analysed in this thesis.

At a more direct level I would like to thank Mr Kenneth Simpson and Miss Liz Hay for their efforts in editing my work.

Without my friends here at the School and the St. John in the City Church I could not have achieved anything since they have generously given me so many opportunities and so much support.

Last but not least, my deepest appreciation is for my family, my wife, Adriana Butter, and my son, Samuel Murwanto, who have made my academic journey here in Wellington full with unforgettable memories, and the new kid in the family, Jonathan Murwanto.

**This thesis is dedicated to my late father Mr Darmo Suwito whom taught me that education was the window to the future.**

# Table of Contents

<b><u>ABSTRACT</u></b>	<b><u>II</u></b>
<b><u>ACKNOWLEDGMENT</u></b>	<b><u>IV</u></b>
<b><u>TABLE OF CONTENTS</u></b>	<b><u>VI</u></b>
<b><u>LIST OF FIGURES</u></b>	<b><u>IX</u></b>
<b><u>I. INTRODUCTION</u></b>	<b><u>1</u></b>
<b>A. RESEARCH PROBLEM</b>	<b>1</b>
<b>B. NEED AND IMPORTANCE OF THE STUDY</b>	<b>3</b>
<b>C. RESEARCH QUESTION</b>	<b>4</b>
<b>D. SCOPE AND LIMITATION OF THE STUDY</b>	<b>5</b>
EVENTS	5
DOCUMENTS	6
ENTITIES	7
SAMPLE	7
<b>E. PREVIEW OF THE STUDY</b>	<b>7</b>
<b><u>II. BACKGROUND</u></b>	<b><u>10</u></b>
<b>A. NEW PUBLIC MANAGEMENT AND PERFORMANCE REPORTING IN THE PUBLIC SECTOR</b>	<b>10</b>
<b>B. THEORETICAL FRAMEWORK AND PRINCIPLES UNDERPINNING PUBLIC SECTOR REFORM IN NEW ZEALAND</b>	<b>12</b>
<b>C. THE ORIGINAL NEW ZEALAND MODEL: OUTPUT-FOCUSED PERFORMANCE REPORTING</b>	<b>15</b>
<b>D. EVALUATION OF THE MODEL</b>	<b>20</b>
<b>E. THE FIRST MODIFICATION: STRATEGIC MANAGEMENT FRAMEWORK</b>	<b>24</b>
<b>F. THE SECOND MODIFICATION: MANAGING FOR OUTCOMES</b>	<b>29</b>
<b>G. AN OVERVIEW OF MANAGING FOR OUTCOMES</b>	<b>33</b>
<b>H. THE CURRENT STATE OF PERFORMANCE REPORTING IN NEW ZEALAND</b>	<b>35</b>
<b><u>III. LITERATURE REVIEW</u></b>	<b><u>38</u></b>
<b>A. THEORETICAL FRAMEWORK</b>	<b>38</b>
1. PUBLIC SECTOR PROBLEMS IN THE PERSPECTIVE OF PUBLIC CHOICE THEORY	38
2. AGENCY THEORY AND THE PUBLIC SECTOR SETTING	45
3. PUBLIC SECTOR REFORM IN THE PERSPECTIVE OF PUBLIC CHOICE THEORY AND AGENCY THEORY	56
4. PERFORMANCE, PERFORMANCE INFORMATION AND THE USERS OF PERFORMANCE INFORMATION	60
<b>B. REVIEW OF PRIOR RESEARCH</b>	<b>63</b>
1. RESEARCH IN PUBLIC SECTOR PERFORMANCE REPORTING	63
2. LITERATURE REVIEW ON DISCLOSURE MEASUREMENT METHODOLOGY AND RESEARCH	67
<b><u>IV. METHODOLOGY</u></b>	<b><u>78</u></b>

<b>A.</b>	<b>OVERVIEW</b>	<b>78</b>
<b>B.</b>	<b>RESEARCH METHOD</b>	<b>78</b>
1.	DISCLOSURE ANALYSIS	78
2.	CASE STUDY FOR ANALYSIS OF RELEVANT FACTORS	86
<b>V.</b>	<b><u>FINDINGS AND ANALYSIS</u></b>	<b><u>91</u></b>
<b>A.</b>	<b>FINDINGS AND ANALYSIS FROM DISCLOSURE ANALYSIS</b>	<b>91</b>
1.	OUTPUT INFORMATION	91
2.	OUTCOME INFORMATION	97
<b>B.</b>	<b>FINDINGS FROM INTERVIEWS</b>	<b>105</b>
1.	MFO INITIATIVE, ITS INTENTION AND IMPACT ON PERFORMANCE INFORMATION	106
2.	STAKEHOLDERS' INTEREST IN PERFORMANCE INFORMATION	107
3.	PERFORMANCE INFORMATION AND ORGANISATION ACHIEVEMENT	112
4.	DEVELOPMENT OF OUTPUT MEASURES	113
5.	DEVELOPMENT OF OUTCOME MEASURES	116
6.	INFORMATION SYSTEM FOR PERFORMANCE INFORMATION	118
7.	MOTIVES FOR PROVIDING PERFORMANCE INFORMATION	120
8.	DISINCENTIVES FOR REVEALING OR IMPROVING PERFORMANCE INFORMATION	123
9.	MAJOR OBSTACLES, CONSTRAINTS AND DIFFICULTIES	125
10.	PERFORMANCE-BASED COMPENSATION	131
<b>C.</b>	<b>DISCUSSION</b>	<b>131</b>
1.	PUBLIC CHOICE PROBLEMS IN THE NEW ZEALAND PUBLIC SECTOR: AN ANALYSIS OF THE POLITICAL CONSTITUTION OF THE WESTMINSTER SYSTEM	132
2.	MAPPING OF THE PRINCIPAL-AGENT RELATIONSHIP IN THE CENTRAL GOVERNMENT SETTING	138
3.	MANAGING FOR OUTCOMES EVALUATION	143
4.	REPORTING INCENTIVES IN THE NEW ZEALAND PUBLIC SECTOR SETTING	150
5.	CHALLENGES IN NEW ZEALAND PUBLIC SECTOR PERFORMANCE REPORTING	154
<b>VI.</b>	<b><u>SUMMARY, CONCLUSION &amp; RECOMMENDATIONS FOR FUTURE RESEARCH</u></b>	<b><u>164</u></b>
	<b><u>REFERENCE</u></b>	<b><u>174</u></b>
	<b><u>APPENDIX A</u></b>	<b><u>186</u></b>
	<b>SELECTED RESEARCH ON PUBLIC SECTOR PERFORMANCE INFORMATION</b>	<b>186</b>
	<b><u>APPENDIX B</u></b>	<b><u>189</u></b>
	<b>SELECTED ACCOUNTING RESEARCH USING A DISCLOSURE INDEX</b>	<b>189</b>
	<b><u>APPENDIX C: LIST OF ITEMS IN THE DISCLOSURE INDEX</u></b>	<b><u>191</u></b>
	<b>OUTPUT INFORMATION</b>	<b>191</b>
	<b>OUTCOME INFORMATION</b>	<b>192</b>
	<b><u>APPENDIX D:</u></b>	<b><u>193</u></b>

<b>LIST OF DEPARTMENTS SELECTED FOR DISCLOSURE ANALYSIS</b>	<b>193</b>
<b>LIST OF DEPARTMENTS THAT ARE EXCLUDED DUE TO UNAVAILABILITY OF ACCOUNTABILITY DOCUMENTS FOR THE PERIOD OF OBSERVATION</b>	<b>193</b>
<b><u>APPENDIX E</u></b>	<b><u>194</u></b>
<b>TABLE 1 RELATIVE SCORES FOR OUTPUT INFORMATION</b>	<b>194</b>
<b>TABLE 2 AVERAGE SCORE FOR EACH DISCLOSURE ITEM – OUTPUT INFORMATION</b>	<b>195</b>
<b>TABLE 3 RELATIVE SCORES FOR OUTCOME INFORMATION</b>	<b>196</b>
<b>TABLE 4 AVERAGE SCORE FOR EACH DISCLOSURE ITEM – OUTPUT INFORMATION</b>	<b>197</b>
<b><u>APPENDIX F</u></b>	<b><u>198</u></b>
<b>DISCLOSURE INDEX RESEARCH INSTRUMENT - SCORING INSTRUCTION</b>	<b>198</b>
<b><u>APPENDIX G</u></b>	<b><u>209</u></b>
<b>SCORING EXAMPLE</b>	<b>209</b>
<b><u>APPENDIX H</u></b>	<b><u>212</u></b>
<b>ETHICS INFORMATION SHEETS</b>	<b>212</b>
<b><u>CONSENT TO PARTICIPATION IN RESEARCH</u></b>	<b><u>216</u></b>
<b>INTERVIEW RESEARCH INSTRUMENT</b>	<b>217</b>



## List of Figures

Figure 1 Accountability Arrangements in the New Zealand Public Sector	17
Figure 2 Principal-agent relationships and accountability arrangements in the New Zealand Public Sector	139

This page is blank.

# **I. INTRODUCTION**

Performance reporting has been one of the most important developments in public sector accounting under the umbrella of the New Public Management movement. In the case of New Zealand public sector reform, the performance management and accountability framework is one of the building blocks of the new public sector management system. The focus of the framework in the New Zealand public sector setting was initially on the performance of output delivery. This focus has been broadened to include the presentation of outcome information. Even though this practice has been going on for more than ten years, there is limited evidence on whether the new reporting regime has improved the quality of performance information disclosed by agencies in the public sector. Furthermore, recent research on performance reporting in central government agencies is very limited. This thesis aims to fill this void.

## **A. Research Problem**

The New Zealand model of public sector reform has generated considerable interest from the academic community since it was introduced in the late 1980s. The model is quite attractive to researchers, due to the pervasiveness of the reform and its strong adherence to underlying economic theories such as agency theory, public choice theory and transaction cost theory. The strong influence of these economic theories has resulted in the adoption of output-focused accountability and contracting in the New Zealand public sector. The New Zealand model was constantly discussed in the public sector literature during the 1990s. However, interest has diminished in the past decade even though the model has evolved from its original version. The basic accountability framework based on output delivery is still in place but it has been complemented by an accountability in the form of management for outcomes.

Public sector performance reporting in New Zealand has formally shifted its focus from solely output-based reporting to a combination of output and outcome reporting following the introduction of the Managing for Outcomes (MfO) initiative in December 2001. This development is seen as an answer to a series of comments on the deficiencies of output reporting (see, for example, Gregory, 2003; Neale & Anderson, 2000; Schick, 1996) and several high-profile failures in the public sector (such as the Cave Creek tragedy and health care sector issues). The aim of this new scheme of

reporting is to enable agencies to provide better performance information to strengthen decision making, to improve communication and interactions with stakeholders and to provide better accountability (Steering Group for the Managing for Outcomes Roll-out 2004/05, 2003). The underlying idea is that better performance information would be provided if output information were supplemented by outcome information.

There is a concern that public sector agencies in New Zealand have too much discretion in reporting performance information under the current reporting regime. Even though the reporting of non-financial performance is mandatory, as regulated in the Public Finance Act 1989, there is virtually no (authoritative) standard to be followed in reporting such information. Therefore, there is a concern that information is presented fairly. There is no direct pressure for public sector institutions to provide information that is useful, except for a set of information that is expected to be reported as prescribed in some guidance documents. There is a concern that the disclosure of performance information is directed more to public relations purposes rather than for accountability purposes. This concern has been confirmed in the recent Beckett and Gilmore report on the outcome of the Treasury's Review of Accountability Documents project (Beckett & Gilmore, 2007).

The state of reporting regulation for non-financial performance information is not as advanced as that governing the reporting of financial information. There are sources of guidance (such as the TPA-9, the Controller and Auditor General's framework of public sector performance reporting and the Treasury's guidance) but none is authoritative. The audit function performed by the Office of Auditor General (also known as the Controller and Auditor General), as reported in its yearly Parliamentary Paper, is directed towards assessment of the information system and control environment for service performance management rather than actual reporting. Furthermore, the audit only provides an assurance on the quality of the performance information reported in the Statement of Service Performance. No review is undertaken of performance information disclosed in other corporate documents, such as the performance planning documents (the Department Forecast Report, later replaced by the Statement of Intent).

The initial emphasis of the MfO initiative was to improve the quality of performance information found in the planning and reporting documents issued by public sector agencies in New Zealand. There was an expectation that this initiative would create an environment in which agencies would be able to provide high quality information in their formal reporting documents. The improvement in the quality of

performance information was seen as a sign of progress from this initiative. However, it is questionable that the desired improvement in the information quality can be realised when agencies can exercise wide discretion in reporting the information and no other governance mechanism has been put in place to ensure that departments provide better information.

## **B. Need and Importance of the Study**

*The only thing that does not change is change itself (Leif Ericsson Leo Veness)*

Reform is about change. Even though reform does not necessarily change the fundamentals of the system, it certainly modifies the existing system. However, it seems that reform has been interpreted as continuous change. The New Zealand public sector system has been continuously changing since the late 1980s when the first ‘big bang’ of change was initiated. The public sector system has not reached a steady state even after two decades since reforms originally started. Adjustments have continuously been made to the system to mitigate aspects deemed to be undesirable. The only certain thing in the reform process has been change itself. It has kept happening.

It has been a continuing challenge in New Zealand public sector reform on how to properly incorporate a focus on outcomes into the system. While the output focus is well established, the public sector system is dysfunctional while it lacks a proper focus on outcomes. Ministers have not been able to formulate adequate outcomes for agencies within their ministries. Governments tried unsuccessfully, using a strategic management framework, to develop some outcome measures. And lastly, government agencies have been forced to provide outcome information, even though they are not responsible for achieving outcomes. This situation seems to have gradually disappeared from public attention. There may be change in the future that will result in proper attention being given to outcomes.

While the initial changes were rigorously investigated by many researchers, subsequent changes have received limited attention. The MfO, which was the last attempt to provide a means by which outcome information in the New Zealand public sector could be provided, has scarcely been reviewed, and then mostly by practitioners. There has been limited attention from academics on this initiative. Most of the reviews have been focused on practical matters with theoretical perspectives being largely absent from the reviews. Furthermore, none have tried to analyse the public sector reforms using the theoretical foundation that underlies the reform. If there are issues

with the reform, before the blame can be directed at incompatibility with the theory applied, it is necessary to first investigate whether the reforms have faithfully followed what is proposed by the underlying theory. This kind of study has rarely been carried out in the public sector setting.

One of the central issues to be addressed in public sector reform is the availability of better information to support public decision making. To be able to investigate this issue, there is a need to develop a measurement technique to assess the quality of information. While there are many proxies for disclosure quality, only simple proxies, such as indicator counts and categories, are used in the current literature investigating the quality of non-financial performance reporting. A more sophisticated technique called a 'disclosure index', commonly used in studies investigating accounting narrative, is adopted only in public sector research that focuses on annual reports. There is a need for a specific disclosure index that can be used in any study investigating non-financial performance information disclosed in the accountability reports of public sector agencies.

There is also a need to get more comprehensive inputs from various actors in the public sector regarding issues concerning public sector reporting. While some reviews offer explanations on the performance reporting issue, there is a concern that these reviews may provide a partial view since all of them were carried out by or on behalf of certain government institutions. There is a limited voice from the preparers of the information in these reviews. There is no review where the various parties involved can give their opinions so that arguments from all important sides are gathered and analysed together. This study attempts to become such a vehicle.

### **C. Research Question**

This study focuses on the issues in the disclosure of non-financial performance reporting by addressing three concerns. The first concern is the level of quality of output and outcome information produced by New Zealand central government agencies. The second concern is whether the latest attempt to improve the quality of performance information, the MfO initiative, has been successfully implemented to produce better quality of information. The third concern is the need to identify the incentives and obstacles faced by public sector agencies in their efforts to produce better output and outcome information for accountability purposes and for public decision making.

The series of research issues and questions addressed in this thesis are:

- whether the current assertions of the disappointing quality of performance information reported by public sector entities in New Zealand are justified;
- how to measure the quality of performance information reported in the corporate documents of public sector agencies;
- since one of the objectives of new initiatives is to improve the public sector system is to address the problem of information quality, a question to consider is the impact of the MfO initiative on the reporting practice of public sector agencies;
- whether there has been progress in the reporting of outcome information by NZ central government agencies over time as a result of the MfO initiative;
- the extent to which the disclosure of performance information by NZ central government agencies has followed the available guidance or benchmark over time; how much effort is required by agencies to meet the reporting benchmark;
- what external and internal factors have driven improvements in the quality of information produced;
- how various actors interact in performance reporting practice.

#### **D. Scope and Limitation of the Study**

The scope of study has been limited by a number of factors, including:

##### **Events**

There were many events that led to the changes in the performance reporting regime. Not long after the original reform in the late 1980s, the government initiated a modification of the public management system by incorporating a strategic management framework. This event led to the introduction of some measures of outcome information in the management and reporting system. There were also several individual initiatives by some public sector agencies to develop outcome focused management, such as initiatives by the Department of Corrections in its pioneering efforts to come up with some outcome measures for its operations. There were also two pilot projects, the Capability, Accountability and Performance (CAP) project and the Pathfinder project, which established the foundation for the implementation of the MfO initiative.

This study focuses only on the impact of the MfO initiative for two reasons. First, there are data availability issues with the other events. Investigating older events forces reliance mainly on document study since it is quite difficult to complement that

with interview or other methods of obtaining oral information. Most of the people who were involved in the processes related to the older events are no longer available for contact. In addition, there is also a problem of getting detailed information of the distant past based on people's memories. Second, some agencies were affected by the other events in particular localities only, therefore the general picture of reporting practice in the government agencies may not have been affected. This study focuses only on events and practice in central government agencies.

## **Documents**

New Zealand government agencies produce many documents that contain performance information. Estimates that are produced yearly contain performance information at output class level and some outcome information. An Output Plan, which is a yearly contract between a minister and the department/ministry, contains detailed performance information of output to be delivered. The Statement of Intent (formerly known as the Departmental Forecast Report) contains a shorter version of information in the Output Plan and some outcome information, in addition to projected financial results. The annual report contains performance information of delivered outputs, some outcome information, and the financial results of the ministry/departments. Some agencies may produce a Strategic Plan document that may contain some performance information. And lastly, there is a three or five yearly report issued by some government agencies that contains outcome information on certain issues, such as Social Report, Environmental Report, or Defence Long Term Development Plan.

This study is limited to investigating performance information in the Statement of Intent (SOI) and the annual report, for two reasons. First, these two documents serve as a pair to fulfil the public accountability requirements of government agencies. The SOI provides the targeted performance information. The annual report discloses the actual results of performance. Meaningful performance information should contain both the targeted and actual indicators. Since the information in the SOI can be paired with the relevant information in the annual report, both provide meaningful performance information. Other documents mostly provide targeted measures, without any indicators showing the results, such as Estimates, Output Plans and Strategic Plan documents. Other documents provide only actual indicators of performance such as a Social Report or Environmental Report. Second, only the Statement of Intent and the annual report are affected by the MfO initiative. Since this study is focused on this initiative, logically the documents targeted for analysis should also be the ones affected by this initiative.



## **Entities**

There are many entities in the New Zealand public sector, such as departments, ministries, Crown entities, local government, schools, boards, etc. This study limits the analysis of performance reporting practice to that of New Zealand central government agencies. This limitation is necessary due to resource and time constraints. In addition, there is a concern that targeting different types of institutions may make the investigation too complex since such a study would have to incorporate differences in the characteristics of the institutions and in their accountability arrangements. Therefore, the validity of the findings of this study is confined to New Zealand government ministries/departments only.

## **Sample**

Due to data availability issue, this study could not target all agencies. It focused on agencies in which all sets of the SOI and annual report during the period of 2003/2004 to 2007/2008 were available. Furthermore, for the purpose of testing the reliability of its research instrument, tests were conducted on a sampling basis only, due to cost concerns. Thus, the validity of the findings may be restricted. Later, in case studies, a sample of nine ministries/departments was selected in order to analyse their reporting practices. However, this does not mean that the findings cannot be used in developing hypotheses or in carrying out further studies.

## **E. Preview of the Study**

This thesis is organised into six chapters. The description of each chapter is as follows:

Chapter 1 introduces the study by stating the problems and underlining the need to investigate the problem. There is also a list of the research questions explored and the scope and limitations of the study are outlined.

Chapter 2 describes the background to the study for the purpose of establishing the context in which the problems were investigated. It describes how the New Public Management movement has changed reporting requirements in the public sector. It is followed by a discussion of the theoretical framework and principles of the public sector reform in New Zealand. The discussion continues with a historical review of

performance reporting practice during three regimes of performance reporting: the original model, the first modification with the strategic management framework initiative, and the second modification resulting from the implementation of the MfO initiative. Between consideration of these three regimes, reviews of the evaluation of the original model and the strategic management framework are presented. This chapter concludes with a discussion on current performance reporting practice in New Zealand.

Chapter 3 presents a comprehensive review of the literature relating to the theoretical framework used, the prior research on the topic, and the methodology used. The two theories that used in the analysis are presented in this chapter: public choice theory and agency theory. Public choice theory focuses on public sector problems and provides a rich template for analysing the behaviours of actors in the public sector setting. Agency theory is discussed with special consideration of its application to analysis of issues in the public sector. Later, these two theories are combined in the analysis of public sector reform in general.

Chapter 4 outlines the research methodology for this study. It provides details of development of the disclosure index, reliability and validity testing, and scoring the documents. It also describes the development and the implementation of case studies, including the interview technique used and the analysis of interview results.

Chapter 5 presents the findings. The findings are presented in two parts: the findings from disclosure analysis and the findings from interview. The findings are described in six parts, starting with general discussion and leading into discussion of more specific areas later. The first two parts provide a foundation for the analysis in the last four parts. These two parts consist of a discussion of problems in areas of public choice in the Westminster system, (the political system in New Zealand), and a map of principal-agent relationships in the New Zealand central government setting. The third part provides an evaluation of the MfO initiative. The last two parts present an analysis of the reporting environment of the New Zealand central government, and discuss the factors that may promote or inhibit reporting practice.

Chapter 6 provides the conclusions. It outlines the nature and purpose of the study. It also summarises the research results and the discussion of the findings. Recommendations are put forward for policy makers while suggestions are made for

further research in the area of performance reporting in the context of central government agencies.

## **II. BACKGROUND**

The New Public Management (NPM) movement during the 1980s led to the adoption of non-financial performance reporting as one of the core and visible components of the public management system. More performance information was required, as contracting processes were increasingly used in the management of public sector agencies. The application of explicit standards and measures of performance was necessary for the functioning of the contract system in which one party has to perform activities or deliver results to the other party according to some agreed standards of performance. Consequently, performance information was required to be available at the end of and/or during the contract duration to enable the other party to assess whether performance or delivery was in accordance with the contract. The framework for reporting performance information has also been a focus of public sector reform in New Zealand. In order to provide an appropriate perspective for this research, it is important in this part of the study to provide background information on performance reporting practice in relation to NPM and the evolution of performance reporting practice in the New Zealand setting. Changes in the reporting framework are highlighted and several potential issues related to the research situation are identified and discussed.

### **A. New Public Management and Performance Reporting in the Public Sector**

Researchers coined the term New Public Management to describe the adoption of private sector management practice in public sector reforms during the last three decades. Glynn and Perkins (1997) argue that NPM philosophy has three principal tenets: the adoption of private sector management concepts and styles, the introduction of quasi-markets and contracting processes, and the application of explicit standards and measures of performance. There was a strong conviction that the private sector is more efficient and effective, and therefore it seemed necessary to adopt the good practices used in private sector management to improve the efficiency and effectiveness of public sector management. The presence of the market facilitated the formation of price, which in combination with contracts was seen as a powerful mechanism for resource allocation. In addition, when explicit standards and measures of performance were more directed towards output rather than input, they were regarded as giving more flexibility to public sector managers to deliver effective and efficient services.

Dunleavy et al. (2006) argue that NPM combines the three approaches of disaggregation, competition and incentivisation to improve management of the public sector. As results and efficiency are rigorously pursued, government institutions are split into smaller and more fragmented units that only deliver products/services with similar characteristics. Agencies deal with each other on a user-pays basis. Competition is fostered between different public agencies and between public agencies and private institutions using quasi-markets. Public managers are rewarded on the basis of their performance in delivering the products/services and managing the agencies. Consequently, the management style in NPM philosophy emphasises output targets, limited term contracts, monetary incentives and freedom to manage (Bevir, Rhodes, & Weller, 2003).

The definition, measurement and reporting of performance is an important element of the NPM. The specification of performance, in terms of output and/or outcome, for the accountability process is the building block of public sector reform around the world. These performance indicators, which are oriented to accountability in terms of results, are seen as a part of the NPM (Hood, 1995). As Behn (1995) identified, the performance management issue is one of the three big questions in contemporary public management.

The form that accountability and performance management takes is defined by how the term performance is further specified. Two foci are commonly referred to in the performance specification: outputs and outcomes. Output refers to the goods and services delivered by government agencies. Outcomes are impacts on the communities caused by the delivery of outputs. Reformers of public sector management would focus accountability on outputs, outcomes, or a combination of both. This focus is not always fixed all the time. As in the case of the New Zealand model, there has been a strong pressure to shift the focus of performance reporting from an output orientation toward a combination of output and outcome information in the last two decades.

The use of performance information in public sector management depended on the concept of accountability adopted in the reform. Any reform that views accountability as control will use contract-based performance reporting (Mayston, 1993). First, the performance aspects that should be accounted for by public sector managers are specified. Once the performance aspects are defined, performance indicators are selected for each aspect of performance. Targeted indicators are then presented in the planning documents of government agencies. At the end of the period, agencies report on the achievement of the performance targets. In this way, the planning

documents serve as a contract of performance, fulfilment of which is reported on in the annual report. The audit function is carried out to ensure that performance information is fairly reported.

Another view of accountability is as a means of dialogue (Roberts, 2002). In this setting, the performance information that is reported allows for a dialogue to take place between the agency and its stakeholders and this can follow disclosure-based performance reporting. For example, outcome information in New Zealand government reporting would be presented as additional disclosure, to accompany the financial statements and the statement of service performance. There is limited audit work on this information, the main aim being to ensure there is which is mainly to assess its consistency with the financial information.

An orientation toward performance indicators is not without criticism. Smith (1995) lists three unintended effects of this situation: ossification, tunnel vision, and suboptimisation. Too much focus on performance indicators can inhibit innovation, especially if accountability of performance is the main focus in performance reporting. A lack of innovation can lead to the problem of ossification. An organisation is then incapable of producing actions beyond its perimeter. The problem of tunnel vision occurs when managers focus on an area that is quantified in the performance management scheme at the expense of qualitative aspects of performance. Similarly, a suboptimisation problem will occur when managers focus on the achievement of their narrow local objectives at the expense of the objectives of the organisation as a whole.

## **B. Theoretical Framework and Principles Underpinning Public Sector Reform in New Zealand**

The New Zealand model of public sector reform is strongly influenced by three economic theories: agency theory, public choice theory and transaction cost theory. Officials from central agencies<sup>1</sup> initiated the reform in New Zealand after issuing a series of policy papers that discussed a strategy to rectify the perceived deficiencies of New Zealand public sector management. When the concepts and principles of the NPM were adopted in designing the New Zealand model of reform, these theories served as the theoretical framework for the model. Boston (1992) argues that the influence of

---

<sup>1</sup> The term “central agencies” refers to three agencies in the New Zealand state sector that provide leadership within the state sector. They consist of the Treasury, the Department of the Prime Minister and Cabinet and the State Services Commission. They have a state sector system-wide perspective, and engagements and connection with every other organisation in three important areas: government strategy, budget, and performance.

these economic theories features in the three most important features of the New Zealand model. First, the principal-agent relationship in the New Zealand public sector is clarified so that agents will not face situations in which they have to serve multiple (and possibly conflicting) principals. Second, the way the incentive structure is designed is based on this clarified relationship and is intended to induce the agents to act in accordance with the interests of their principals, which will mitigate the public choice problems of the bureaucrats or the agency problems from the perspective of agency theory. Third, changes are made to ensure that the transaction cost is minimised in institutional arrangements by examining the relative costs and benefits of employing and supervising agents versus the option to purchase externally.

In public policy and management, two broad classes of problems have emerged from the transaction cost perspective: public choice problems and agency problems. The public choice problems stem from the capture of policy and resources by organised and better-informed special interests as well as public institutions. There are two aspects of public choice problems: the first is related to political decisions and the other deals with bureaucratic decisions. In political decisions, politicians may decide to fund programmes or projects directed to benefit their local and/or specific constituents instead of ones that benefit the majority of voters. In bureaucratic decisions, concerns are raised that public sector bureaucrats may not act in the best interests of their political principals. This is essentially an agency problem, which is caused by employing people in public delivery systems whose incentives need to be aligned with the legislature's or governing bodies' established policy objectives. This alignment results in a class of transaction costs known as agency costs. These costs include the costs of establishing, monitoring and enforcing contracts for service delivery and deadweight losses due to contract imperfections.

The aim of the New Zealand model, as stated in the Steering Group's review of the reform in 1991, shows the influence of these three economic theories. The reform is aimed at establishing a policy framework based on 'Clear, prior specification of intended performance, appropriate delegation of decision-making authority, subsequent monitoring of achievement and the careful application of incentives and sanctions' (Steering Group, 1991, p. 60). From the public choice perspective, essentially this is to control the bureaucrats so that their capture of government policies is minimised. From the agency theory perspective, similar controls over bureaucrats are needed to minimise agency costs.

Bale & Dale (1998) argue that there are five characteristics embodied in the New Zealand model that clearly show that the model is strongly based on ideas borrowed from public choice theory (especially in mitigating the public choice problems which stem from bureaucratic decisions) and agency theory. The characteristics are:

- Clear lines of accountability are established between government ministers as the principal and their departments as the agent.
- Department performance is defined in an unambiguous and measurable way. The non-financial performance, for example, is measured and assessed at the output level of the departments.
- Authority is delegated to chief executives, especially in relation to the management of inputs allocated to produce the targeted performance.
- Incentive systems are designed to reward or punish chief executives in relation to achievement of the targeted performance.
- Structures for reporting and monitoring performance of chief executives are put in place.

The Treasury, in its official briefing document to the incoming government in 1987, as quoted by Cook (2004), formulated five cascading principles underpinning the New Zealand public management system: clarity of objectives, freedom to manage, accountability, effective assessment of performance, and adequate information flow. These principles show the strong influence of public choice theory and agency theory in the New Zealand public sector reform, in which performance assessment, incentives and information play an important role in the relationship between principals and agents. The starting point of a well functioning public management system begins with the specification of objectives. The objectives that chief executives are responsible for achieving should be specified as clearly as possible. If possible, the objectives should be identified individually and multiple and conflicting objectives should be avoided. The next step is the delegation of authority and power to make resource allocation decisions in order to achieve the objectives efficiently and effectively. This condition requires the ability of chief executives to freely manage inputs. Good public sector management should complement this freedom to manage principals with a proper incentive system and accountability mechanism. A system that provides incentives and sanctions to modify the behaviour of chief executives should be in place, so that they pursue the established objectives instead of their own goals. Chief executives must be accountable for their performance and their principals must be able to ask for accountability and to



reward accordingly. The performance should be assessed effectively, and the assessment must be able to differentiate between the controllable and uncontrollable aspects of objective achievement. Lastly, performance assessment needs sufficient quantity and quality of information concerning the chief executive's performance.

### **C. The Original New Zealand Model: Output-Focused Performance Reporting**

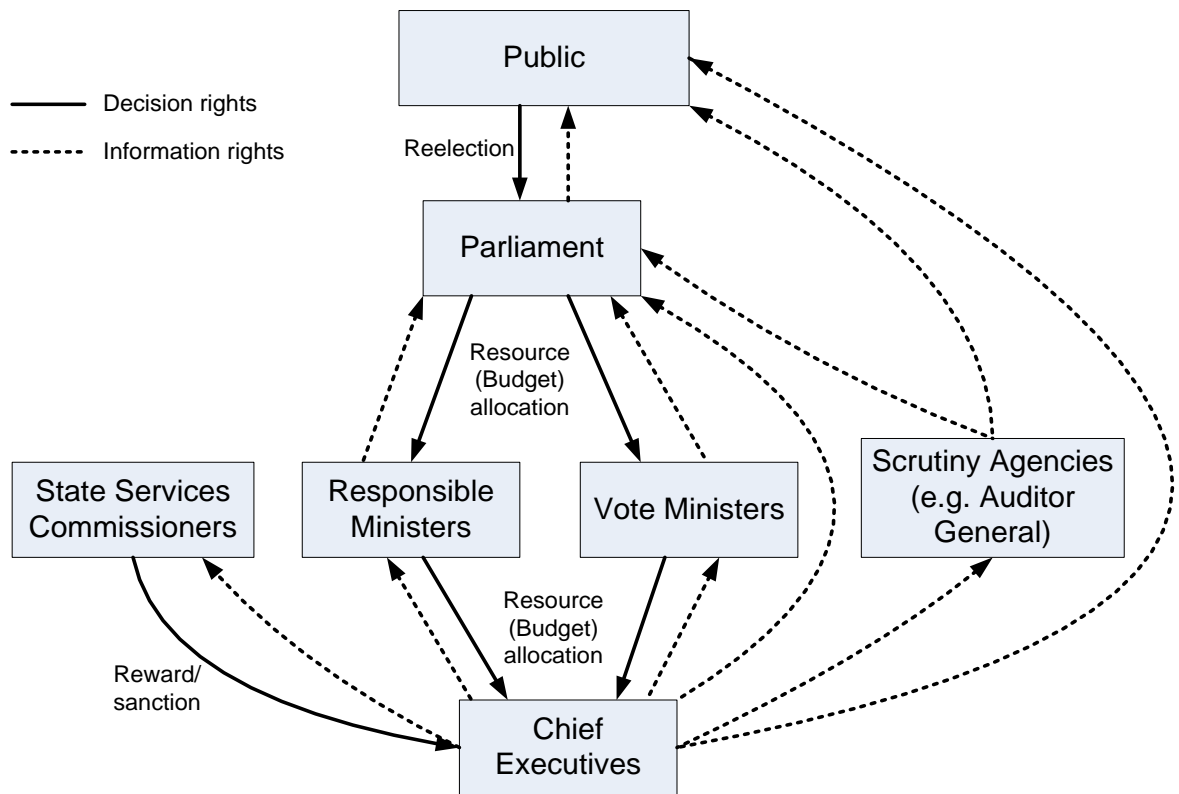
The State Sector Act 1988 (SSA 1988) and the Public Finance Act 1989 (PFA 1989) reflect public sector management from the perspective of agency theory, in which the principals, agents and the relationship between them are modelled, defined and structured following a set of principles (Boston, Martin, Pallot, & Walsh, 1996). The principal for the executive agents is defined very specifically as ministers, replacing the previous arrangement of multiple and ambiguous stakeholders. Chief executives are employed to serve ministers.

There are two critical aspects of the new arrangement. First, the critical link in the accountability chain focuses on the relationship between the ministers and departmental chief executives. The New Zealand model goes one step further by concentrating only on one relationship, while in public sector reforms in other jurisdictions, multiple accountability relationships are still preserved. Second, the tasks of each party in the relationship are set so that the accountability and responsibility of each party are clear (Scott, 2001).

The relationship between the ministers and their agencies is set to have two dimensions: ownership interest and purchasing interest. The ownership interest follows the relationship that exists in the private sector where the principal (government or the Crown via a minister) is viewed as the owner of departments (government agencies). The interest of the owner is similar to the interest of investors in general. They desire the best return possible from the resources allocated to the departments. Public production lowers the wealth of society if the rate of return earned from this investment is lower than the return on comparable investment in private ownership (Scott, Bushnell & Sallee, 1990). As an owner, the principal demands accountability for financial performance from the agent for managing his/her ownership. Financial reporting standards for discharging this accountability are borrowed from the established framework used in the private sector.

The purchase interest, which was developed specifically for the public sector setting, recognises that the same principal has a purchase interest in his/her departments. Principals have the option of acquiring goods and services (outputs) from departments or sourcing them from more competitively priced suppliers in the private sector. To meet the policy objectives of the government of the day, ministers act as a member of the Cabinet to plan and negotiate the services required from their departments. In making their decisions, ministers may ask for advice from the department providing the services and may also employ independent advisors (Scott et al., 1990). The ministers' concern in a purchase decision is to get value for money, as if their department is buying from an independent supplier. The accountability on purchase interest is contractual. At the beginning of the fiscal period, there is a contract that specifies the outputs to be delivered by the departments. This contract is documented in detail in the Purchase/Output Agreement and reported in summary in the Statement of Intent (Department Forecast Report). At the end of the fiscal year, the achievement of the agencies in providing the planned outputs is presented in the Statement of Service Performance (SSP) and is then assessed.

Boston (1992) further elaborates the task of ministers and chief executives in the purchase interest form of the relationship. As principals, ministers are responsible for choosing the outputs to be purchased and for the outcomes associated with these choices. They have the task of setting policy agenda, determining departmental priorities, specifying their desired outcomes, deciding on the outputs necessary to achieve these outcomes, and monitoring the chief executive's performance in delivering the outputs. As agents, chief executives are responsible for delivering the outputs selected. They have the task of ensuring that the minister's requirements are satisfied and of taking responsibility if they cannot deliver the outputs as specified in the contract. The accountability is structured in a staggered scheme, in which executives of the agencies are formally accountable for outputs via contractual arrangements, and the ministers are supposedly held accountable for the outcomes to parliament (the public) via the political arrangement. In addition, executives of the agencies have other forms of accountability to the State Services Commissions and other Scrutiny Agencies (see Figure 1 below).



**Figure 1 Accountability Arrangements in the New Zealand Public Sector (Adapted from Anderson & Dovey, 2003).**

The sole focus of chief executives' accountability to ministers is one notable characteristic of the New Zealand model that differentiates it from public sector reform models in other jurisdictions. In contrast, the models in other jurisdictions incorporate the accountability of heads of government agencies directly to the public who use the public service, in addition to the accountability to ministers. The accountability of chief executives of New Zealand agencies to the users of services is exerted indirectly via ministers and Parliament. Public agencies are accountable to ministers and to Parliament which, in turn, is accountable to the voters who elected the representatives. This accountability arrangement is in line with one of the underlying principles of the reform, which is the clarity of objectives. If managers face uncluttered lines of accountability, they may have a better chance of managing efficiently and effectively. On the other hand, if public managers face twin direct accountability to ministers and members of the public, there is a possibility of accountability conflicts, especially among the front-line staff. They may face a dilemma in their loyalties, of whether to focus more on serving the members of the public or to serve their management objective of serving the minister.

The specification of public services contained in the annual contracts for service provision between ministers and chief executives are based on the upward accountability to ministers and Parliament in terms of the suitability of agreed outputs for the preferred outcomes. The contractual obligation is between ministers and chief executives. The contract specifies members of the public as beneficiaries of service provision, but they are not formal parties to the contract. Therefore, they cannot directly hold chief executives accountable. As Mulgan (2008) argued, this accountability arrangement reflects the purer form of the Westminster system in which the central mechanism of government accountability is the chain of ministerial and parliamentary accountability. Officials answer to their superiors, the chief executives, who then answer to ministers. Subsequently, ministers answer to Parliament and lastly Parliament answers to the people.

In addition to clear specification of objectives, responsibility and performance, it is necessary to devise the right incentive system and establish proper authority. The new arrangements incorporate a new type of incentive by applying labour-market mechanisms. Permanent tenure for departmental heads is abolished and replaced by term appointments. They are called chief executives. The State Services Commission (SSC) manages the appointment process. Each appointment is made on contract for up to five years, with the possibility of reappointment. The contracts also contain terms of employment, which are negotiated by the SSC. The Governor-General in council makes the appointment decision, based on recommendations submitted by the SSC. The Prime Minister and Minister of State Services approve the term of employment. This mechanism, as Scott et al. (2001) argue, ensures that ministers have little ability to influence the identity of chief executives, other than at initial employment, or any of their conditions of employment. This is essential in the Westminster system to ensure that bureaucratic appointees remain politically neutral and only serve the government of the day.

Scott (2001) argues that if properly implemented, the new public management framework promotes a degree of detachment from the past legacy of input-based public sector management in which ministers and central agencies had to argue for the added value of a certain service from a position of less information than the department had. With the identification of ministers' objectives and requirements, there is always a need to question the existing services and their costs, whether they are appropriate for these objectives and requirements. The chief executives cannot assume that the past service will still be required. There are incentives in place for departments to produce outputs

arising from an analysis of current and future objectives and requirements. Chief executives have to demonstrate the value added aspect of their existing and proposed services.

As mentioned by Scott *et. al.* (1997), the most controversial element of the New Zealand reforms was the selection of outputs, rather than outcomes, as the measures of non-financial performance for accountability purposes. Reformers claimed that enforcing accountability to chief executives in terms of outcomes would not be workable as there are problems of measurability, causality, control and time frames. The focus on output accountability in the New Zealand model is in line with the principle of performance measurement which requires such measurement to be carried out in an unambiguous and measurable way, so that the accountability relationships between ministers and executives are clearer and stronger. As ministers are still responsible for outcomes, there is no intention to reduce the attention on outcome achievement. There is an expectation that while government agencies are tightening the management process by understanding the value of the concept of controllable outputs, they are also identifying and pursuing better outcomes (Scott, 2001).

The selection of outputs as the focus of accountability has been based on some perceived problems in the New Zealand public management system. Scott (2001) identifies three problems that were intended to be addressed by output-focused accountability. First, ministers did not always have enough understanding of and/or give enough attention to the range of activities being carried out by departments. They may have given attention to some areas of their interests but they were frequently poorly informed on other areas of departmental services. Ministers also may have had a perception that many of the services in the areas outside their concern were not highly valued, either by themselves or by the recipients of the services. Second, the responsiveness of bureaucracy to changes in policy was still not on a level that ministers considered satisfactory. Government agencies were not fast enough to respond to changes in politicians' policy preferences. Even though ministers had an array of administrative controls at their disposal, they felt that they did not have the degree of control necessary to ensure that their policy preferences were being implemented. Third, there was a perception on behalf of ministers that the accountability mechanisms for departmental heads were very weak. There were no appropriate incentives, as the tenure of departmental heads was permanent. In addition, there was no prior specification of service delivery in a given period so that it was difficult to extract information subsequently in order to provide effective accountability.

## **D. Evaluation of the Model**

After a decade of implementation, the new accountability arrangement has generated mixed results. Researchers have claimed that the reforms have improved accountability in the New Zealand public sector setting. The reforms also brought enormous efficiency gains in the NZ public sector (see for example, Newberry & Pallot, 2004; Schick, 1996; Scott et al., 1997). The flexibility introduced by the accountability arrangements has enabled public managers to do more with fewer resources.

On the other hand, researchers are also quick to raise their concerns about the New Zealand model in several respects. First, a system that focuses on output is not compatible with ownership interest. Second, the accountability for outcomes has not been formalised in the system. Third, coordination among public sector agencies for pursuing shared outcomes is greatly undermined in this model. Fourth, as outcomes are not explicitly stated, there is a concern that the decisions on outputs may not take into account the consideration of outcomes. Lastly, there is an ongoing issue with output specification or output definition making it an uninformative guide to what the public would expect from public agencies.

Some researchers have raised concerns that the focus on outputs and the purchase interest of the principal will undermine the ownership interest of the same principal. Newberry & Pallot (2004) quote the concern raised by the State Services Commission that these arrangements will erode the long-term capability of agencies and thus the ownership interests of government in the agencies, the same concern as previously raised by Schick (1996). This ongoing resource erosion could eventually make departments lose the capability to perform even their core functions (Audit Office, 1999). As Campbell (2001) has put it, there is a serious distortion in the relationship between ministers and public managers as a result of an overemphasis on outputs.

Some other researchers argue that the accountability arrangements of the New Zealand model do not operate properly in practice, especially in relation to the accountability of the ministers. The principals (ministers) are accountable for the outcomes of their agencies under the political accountability arrangements. The Public Finance Act 1989 (PFA 1989) requires ministers to identify the link between the outputs and their (Government's) outcomes in the Estimates submitted to the parliament. However, ministers, for the most part, do no more than assert in the Estimates that certain outputs contribute to certain outcomes. In addition, there is no requirement for ministers to devise measures of performance to demonstrate progress towards

achievement of outcomes (Ussher & Kibblewhite, 2005). When the PFA 1989 was being discussed and considered, the Controller and Auditor-General raised concerns about the lack of requirements in the bill for ministerial reporting on outcomes (Neale & Anderson, 2000). The lack of a reporting requirement on outcome information, and the separation of output and outcome accountability, are key factors in the unsatisfactory linkages between agencies' outputs and the ministers' intended impacts on spending programmes (Webber, 2004). As quoted by Scott et al. (1990), the comment of the Controller and Auditor General in his submission to the Finance and Expenditure Select Committee of Parliament on the Public Finance Bill gives a good summary of this issue. The New Zealand model falls short of being a well-structured set of accountability and reporting provisions because it has a significant omission in the provision for ministerial reporting on outcomes (Finance and Expenditure Committee, 1989, p. 18).

The case of the Cave Creek disaster, and the subsequent enquiry, shows how the accountability arrangements in the New Zealand model were tested as to whether proper accountability would rule when problems arose in the public sector. On 28 April 1995, a scenic viewing platform in Paparoa National Park collapsed, resulting in the death of 14 people. A commission headed by Judge Graeme Noble conducted an enquiry into this disaster and concluded that nobody individually or collectively was responsible for the tragedy. Instead, the system and the culture within the Department of Conservation were to blame. Consequently, there was no immediate resignation from senior officials because of this disaster. The Minister of Conservation asserted that accountability in this case was provided by his public accountability and his determination to ensure that corrective actions were implemented. The minister insisted that he could claim that he fulfilled his duty of obligation as a minister under the doctrine of ministerial responsibility, if he was prepared to answer fully for the events and could address the flawed system and culture that contributed to the tragedy (Gregory, 1998). Eventually, the minister resigned after the media and public pressured him intensively. A new minister was appointed and the State Services Commission conducted a full review of the Department of Conservation.

In analysing this case, Gregory (1998) argues that in a Westminster regime, the accountability flows from the requirement that ministers are answerable to Parliament and the electorate for their own and their department's actions and/or inactions, not from the ability to make them or their chief executives resign from their positions. This doctrine is seen as inadequate, since ministers have actively tried to avoid resignation on several occasions by distancing themselves from political responsibility when their

department's administration has gone wrong. The New Zealand model was designed to ensure that rigorous assessment of the responsibility of ministers and chief executives could be applied so that rewards or sanctions could be given accordingly. Yet, in the case of the Cave Creek disaster, this assessment could not be properly implemented. The argument for ministerial and managerial responsibility still revolved around the ministerial answerability for the problem and the promise of ensuring that shortcomings were corrected.

Gregory (1998) further argues that in this tragedy it was not easy to differentiate something as an output or as an outcome, even though the responsibilities of ministers and chief executives were clear in this regard. It could be argued that a safe viewing platform was the desirable outcome. Therefore, the outputs that were to be purchased were all necessary components to ensure that the platform was safe, including certain types of bolts that should have been purchased and used to install the doomed platform. The Purchase Agreement should show the detailed performance of the output. On the other hand, another argument could be raised that the output purchased by the minister was a safe platform, not just a platform. Therefore, the minister should apply the *caveat emptor* principle. He should be wary of buying a faulty platform. The main question is whether the separation of responsibility embodied in the public sector reform has led public agencies to be concerned more on measuring outputs rather than producing desirable results. Public managers might focus more on the short-term achievement of output delivery and thus undermine the agency's roles and functions to help ministers realise their desired outcomes.

Another criticism is that the New Zealand model placed too much focus on the horizontal relationship between chief executives and their ministers and too little on their relationship with the Cabinet as a whole (Boston, 1992). Many outcomes sought are shared outcomes that require a combination of outputs produced by several departments. Boston (1992) further argues that there is no mechanism that ensures a coordination of activities within the government agencies to pursue these shared outcomes. There is no involvement or oversight by the Prime Minister or Cabinet in the preparation of performance agreements. There is also no provision that requires a chief executive to consider the shared outcomes of the government and coordinate or cooperate with other departments in the development and implementation of policies across the public sector. Therefore, he concluded that there is an impression that the responsibility of the chief executive is to provide outputs for individual ministers rather than the outputs necessary to achieve the collective interests of the government.



Boston & Pallot (1997) argue that the problems of horizontal coordination between government agencies were intensified in the New Zealand model. The reform has shifted public sector management into a more decentralised system. This system, combined with the accountability focus shared between chief executives, contributes to the problem of coordination. Chief executives argued, when interviewed by a review team on the state sector reforms in 1991, that leadership and strategy (for the government's collective interests) were the constitutional responsibilities of the cabinet and ministers, as quoted by Boston and Pallot (1997). Their expectation was that the government would provide or agree on a set of priorities in the form of policy goals or outcomes. Unfortunately, the Cabinet did not develop or articulate them, so individual ministers and their chief executives could not specify the intended outcomes that represent a collective interest of the government. Hence, horizontal coordination was not properly supported.

There is also concern that there was no systematic evaluation process to ensure that the outputs specified and produced were the best combination of outputs to achieve the Government's objectives or outcomes. Most reviews on the New Zealand state sector reforms, including reviews done by Schick (1996), Scott (1996), and the Controller and Auditor General (1999), have all pointed to issues of linkage between outputs and outcomes. Petrie and Webber argue that these issues have two different aspects. The first aspect is the difficulty in relating specific outputs to broad outcomes and the second aspect is the lack of a mechanism to monitor and evaluate outcomes. In the first aspect, the Controller and Auditor General (1999) asserted that government agencies have narrowly interpreted the term "link" used in the PFA 1989 for relating a class of outputs to outcomes. They simply assert that a class of outputs will contribute to an outcome without an elaborate explanation of how it is expected to do so. In the second aspect, the reporting of outcomes is not developed in the formal system, leading to the lack of a mechanism to monitor and evaluate outcomes.

Petrie and Webber (2006) argue that the main problem/weakness is that departments and agencies could not re-evaluate their activities as a part of an annual process of output specification to find out the best way to contribute to the Government's objectives or priorities. Therefore, no process was in place to seek continuous improvement in the alignment of outputs to outcomes. The problems in the New Zealand model are largely with the effectiveness of government interventions, which is the extent to which the desired outcomes are achieved through the delivery of outputs. The effectiveness appears more problematic in three areas in which the level of

difficulty is quite high, in relating outputs to outcomes: policy advice provision, social service delivery (such as outputs in the area of healthcare, education, housing, entitlements, Maori Development, etc.), and regulation activities.

The last issue with the New Zealand model, as Mulgan (2008) puts it, is the problem relating to the definition of outputs or output classes for which departments are accountable. This problem is quite apparent in the area of policy advice delivery. The decision to focus on outputs in the accountability relationship is based on the preference for results that are less ambiguous and more easily specified and measured. However, some outputs remain difficult to specify meaningfully, so that they cannot be used as informative guides of the public's expectations of public servants. If someone is dissatisfied with the performance of an agency, he or she can point to the specification to highlight his/her points of dissatisfaction. In practice, the purchase/output agreements and output statements are rarely used in public debate on government accountability. Mulgan (2008) argues that output information is bypassed in public debate because it is not specific enough to be helpful in nailing down the government agency's obligation, especially in the area of policy advice in which the output classes seem to be incomprehensible and the performance measures are tedious.

## **E. The First Modification: Strategic Management Framework**

During 1993-1994, not long after the implementation of public sector reform, the government (under the leadership of the National Party) initiated corrective action to the New Zealand model by adopting a longer-term Strategic Management Framework. The government introduced strategic measures under a new policy-making framework in the form of strategic result areas (SRAs). Ministers and officials were required to define the strategic nature of their business as related to these SRAs and thus confirm or revise the key result areas (KRAs) developed for each SRA related to them. Boston & Pallot (1997) state that this framework was developed to address three concerns: inadequate specification of government goals or objectives, the lack of governmental strategic vision in the policy and budgetary process, and the lack of proper coordination among agencies within the public sector.

Before the development of SRAs, there was no formal strategic management process in the New Zealand model. There was a requirement in the PFA that the Estimates should provide information on the link between each class of outputs and the Government's desired outcomes for each vote in the government budget. Therefore, ministers were required to be explicit about the outcomes that their departments were

trying to contribute or achieve. However, as previously mentioned, this requirement is narrowly interpreted by providing a simple assertion about the link, instead of elaborating on it, if outcomes were identified in the Estimates. These outcomes were only identified if the information was available formally through high-quality policy advice provision or informal information dissemination (Cook, 2004).

Boston and Pallot (1997) argue that the adoption of KRAs and later the related SRAs was a combination of two unrelated initiatives. The development of SRAs was a top-down process by the National-led government to develop long-term objectives of the government. On the other hand, the development of KRAs was undertaken by the SSC to try and correct some perceived deficiencies in the new public management system adopted in the late 1980s. This initiative was essentially a bottom-up approach.

They state that the top-down process was started when the SRAs framework was introduced following a failed project of improving the government's strategic management by developing a model that encompassed social and economic strategies. This model was perceived as an extensively complex framework with complicated interconnections and was further complicated by the tendency that every department wanted to be considered as strategically relevant. Developing such a complicated model proved to be an overwhelming task for the government. A simple alternative model was developed under the leadership of David Kirk, political advisor to the prime minister at that time, by preparing a policy statement setting out the National Party's long-term objectives for government. This policy statement was published in a document entitled "Path to 2010" about four months before the general election of 1993. The document contained the major objectives of the government, but did not spell out a detailed programme to achieve these objectives. When the National Party won the 1993 election, they faced the problem of translating the broad vision in "Path to 2010" into more detailed policies and a clear set of priorities.

Boston and Pallot (1997) also state that the bottom-up process was started at the same time under the leadership of the SSC and the Department of Prime Minister and Cabinet (DPMC). During the early 1990s, it was becoming apparent that there were some deficiencies in the new public management system for performance specification, monitoring and assessment. The most unsatisfactory aspect was in assessing performance, as the quality of performance agreements varied from one agency to another and they also lacked specificity. The SSC then developed the idea of KRAs as the key measures for which chief executives were held accountable. KRAs were the most significant objectives set for chief executives to achieve during a financial year.

This initiative was essentially an effort to align the interests of the SSC in assessing the performance of chief executives and the interests of chief executives for which their performance was assessed (Boston & Pallot, 1997). The SSC wanted to shift public management from quantitative and output-based management into something more qualitative and focused on outcomes, which was missing in the original model of New Zealand public sector reform. This focus was also longer term to provide equal attention to the ownership interests of government, which by nature tends to be longer term. Chief executives wanted to see the delivery of outputs placed in the context of government priorities.

The DPMC played an important role in the development and implementation of SRAs, since coordination was needed to ensure that SRAs for each individual agency could link to the government goals described in “Path to 2010” (Boston & Pallot, 1997). The term SRA came later in 1994 during an early exercise of the SSC and DPMC in reviewing the performance agreements of chief executives. The DPMC introduced the term SRA to serve as a bridge between the broad objectives of the government in “Path to 2010”. KRAs were introduced into the performance agreement of each agency. Under the DPMC directions, departmental chief executives decided on the SRAs appropriate for their departments in terms of three areas that they had to consider. First, they should find out what their ministers wanted to achieve in their portfolio area. Second, they should identify the two or three objectives that their department needed to meet to achieve the goals set out in “Path to 2010”. Third, they had to ensure that those identified objectives corresponded with the priorities of their ministers. With these directions, the DPMC had asked chief executives to consider the government’s collective interests embodied in “Path to 2010” when framing their departmental outcomes.

The risk management process was also part of this strategic management exercise (Boston & Pallot, 1997). The DPMC required chief executives to identify some of the most serious risks or key issues that might jeopardise the department’s efforts in achieving the government’s collective interests and key policy goals. These identified key issues or risks were later confirmed with the respective ministers when senior DPMC officials met with them individually.

The result of the SRAs development process was a paper that outlined a new strategic management process that government agencies should follow and listed all identified specific SRAs (Boston & Pallot, 1997). The paper was discussed and later endorsed by the Cabinet in May 1994 after it was referred to the Cabinet Strategy

Committee for consideration. Even though the initial set of SRAs was considered as an internal management tool, it was later available for the public in February 1995, before the first budget policy statement under the new Fiscal Responsibility Act was produced. The SRAs document contains a description of each policy area identified, the objectives for each policy area in terms of outcomes and a list of activities to be carried out during the next three to five years to achieve each objective. The document serves as a reference for government agencies in selecting the appropriate SRAs. Since chief executives were involved in the development of SRAs, the objectives were directly applicable and relevant to them. With the SRAs, the strategic objectives of the government are integrated with the performance of departments.

The SRAs were expected to be the basis for yearly plans of government agencies. They served as the outcomes which chief executives had to consider in planning and offering the set of outputs to be purchased by their respective ministers. However, KRAs, instead of output information, were used as a basis for performance assessment since they were viewed as measures that were more specific and less varied. Each year, ministers and chief executives confirm, and revise when necessary, a series of KRAs based on the published SRAs. The KRAs were developed only for areas critical to a department's business. If the activities conducted by departments related to SRAs, chief executives had to specify the particular SRAs to which their departments would contribute. In these cases, KRAs would show an emphasis on the achievement of particular SRAs. If the activities were not related to SRAs, emphases in the KRAs reflected the minister's priorities and the strategic effectiveness of the organisation.

This new development translated into the practice of linking departmental outputs with the identified outcomes (defined as SRAs and measured using KRAs). In planning documents (Departmental Forecast Reports), executive agencies identified the SRAs related to their operations and stated the KRAs. In developing the annual planning documents, which contain output specifications, executive agencies discussed and re-evaluated with their respective ministers how their outputs might best contribute to achievement of the government's objectives (SRAs and the related KRAs). In the annual report, the longer-term outcome accountability is discharged by requiring executive agencies to report on progress towards achievement of the planned KRAs. This process is expected to enable executive agencies, together with their ministers, to seek continuous improvement in the alignment of outputs and outcomes.

With the introduction of KRAs into the New Zealand model, chief executives were induced to adopt a longer horizon in their planning. The SSC expected that the

KRAs did not vary greatly for the three-year period corresponding to the election cycle period in New Zealand, unless there was a major change in strategic direction agreed by ministers (Boston & Pallot, 1997). The basis for assessing the performance of chief executives and departments was observable and verifiable milestones were provided for each KRA. Therefore, SSC required chief executives to provide these milestones in detail (including the dates of delivery) so that the progress towards achievement could be measured.

One of the strong points of the introduction of the strategic management framework in the New Zealand model was the identification of shared outcomes and thus the facilitation of coordination among relevant agencies in developing strategies for achieving these outcomes. There were two mechanisms for promoting the identification of shared outcomes and the coordination of activities among agencies. First, in the performance agreement, there was now a requirement for chief executives to cooperate and consult with other departments where appropriate (Boston, 1992). Central agencies monitored the fulfilment of this requirement and included the results as a part of the annual performance assessment process. Second, the State Services Commissioner reviewed each draft of the KRAs with the assistance of the Secretary of Treasury and the head of the DPMC before chief executives and ministers could finalise and sign the performance. The review assessed how good the specification of KRAs was, whether they were internally consistent, whether they were congruent with the KRAs of other departments and how compatible the KRAs were in contributing to a shared SRA (Boston & Pallot, 1997). This review also ensured that all relevant parties had been consulted in the KRAs development.

However, this modification did not produce the intended result of mitigating the concerns held about the original New Zealand model. Various reviews have been conducted and so far, the reported results have been mixed. The benefits mentioned are better clarification of the Government's strategic objectives, improvement in the "conversation between ministers and chief executives" regarding strategic planning (State Services Commission, 1998a), and better specification of outputs (Petrie & Webber, 2006). The weaknesses identified related to the goals (SRAs), the measures (KRAs) and the assertions of links between outputs and goals. The goals are too broad to be useful as strategic directions (Ussher & Kibblewhite, 2005). On the other hand, the specification of KRAs has been seen as too narrow (Controller and Auditor General, 1999; Petrie & Webber, 2006) and there have been a lack of targets or quantifiable measures of progress over time (Ussher & Kibblewhite, 2005). Many KRAs also

involved milestones which lacked challenge (Controller and Auditor General, 1999). In addition, the relationship between outputs and goals has been interpreted narrowly by simply asserting that a class of outputs will contribute to an outcome without describing how it might be expected to do so (Controller and Auditor General, 1999). The SSC (1998b) also concluded, as quoted by Petrie & Webber (2006), that the main area of concern is a weak connection between strategy and spending. They questioned the appropriateness of the application of private sector approaches to strategic management in public sector institutions. The strategic management approaches of the private sector are applied in an environment where there is a single coherent sense of purpose, whereas government institutions typically have multiple objectives that cannot be coherently combined. The strategic management approach cannot be adopted easily without making appropriate modifications.

Although the Strategic Management Framework had been successful in some respects in drawing the attention of executive agencies to the importance of the relationship between their activities and the expected or intended policy impacts (Webber, 2004), this initiative had been largely abandoned by the late 1990s. Webber (2004) argues that political and bureaucratic factors were behind the reasons for this abandonment. The political factor could be seen in the way in which the new Labour-led coalition government was determined to roll back some of the policy reforms of the 1980s after the coalition won the election in 1999. In relation to the bureaucratic factor, there was a growing perception that executive agencies did not give sufficient attention to the results of their activities in their accountability for outputs. The KRA initiative was unable to remedy this situation since it is not a formal correction of the perceived shortcomings in the government's outcome specification and the links between outcomes and outputs (Controller and Auditor General, 1999). The Auditor General (1999) also noted that there is a risk that this framework is unlawful, as it is not a part of any legislation.

## **F. The Second Modification: Managing for Outcomes**

The origin of Managing for Outcomes, the next adjustment to the New Zealand model of public management, can be traced back to the analysis and recommendation of the Review of Centre (RoC) published in 2001 (Webber, 2004). The RoC is a review conducted by a committee of officials from central agencies with input from a small number of external advisers. In the review process, they broadened their focus from just reviewing the roles and activities of the central agencies into assessing the broader

public management issues, particularly analysing various shortcomings that had been identified in the original New Zealand model. The Advisory Group on the Review of the Centre (2001) argues that there are three main categories of perceived problems in the model. First, there is a lack of integration in the delivery of government services. Second, the activities of departments and agencies within and between sectors are fragmented because there is a lack of an alignment mechanism. Third, the building and sustaining of institutional capabilities and organisational culture is considered weak. The RoC's recommendations are focused on incorporating a much stronger outcomes focus into the budgeting and expenditure management process to complement the existing output-based model.

One thing that is similar between the strategic management framework modification and the MfO modification is that they are both an incremental approach to the existing system. Both modifications are intended to complement the existing output-based system. The strategic management framework was aimed at combining the top-down strategic management process with the bottom-up output-based planning. The RoC considered that the MfO is aimed at softening the output-based model rather than replacing it with a new or different approach. The focus of the MfO-based public management system is to incorporate the balanced perspectives of outputs and outcomes. While ministers are still responsible for outcomes, chief executives are now responsible for "managing for outcomes", in addition to their existing responsibility of delivering the agreed outputs. The desire to keep the existing output-based system intact was supported by the government of the day when the Minister of Finance in 2003 stated that the output-based model was viewed as sound, so there was no need to make a major amendment (Webber, 2004).

The difference between the two modifications is in the approach to outcome identification and reporting. The combination of top-down and bottom-up approaches in the strategic management framework led to the establishment of a flawed system in which the development of output and outcome measures was not integrated. The end result was an attempt to combine three incompatible mechanisms (the "Path to 2010", the KRAs and the output-based management) into a single strategic planning system that seeks to align the internal strategic planning within government agencies with that of the government. The MfO can be considered as a bottom-up approach in which the departments play key roles in identifying, measuring and reporting outcomes, by consulting with their respective ministers. The internal strategic planning process should ensure that the strategies developed are aligned with the outcomes. Therefore, the RoC



recommended that there should be a requirement for the departments to develop procedures and documents in their annual planning activities to strengthen their internal strategic planning process, to ensure that outcomes are identified, measured and monitored.

At the same time the Review was conducted, the central agencies implemented two pilot projects to improve public management with two objectives: assessing the viability of outcome-based management being implemented in the New Zealand public sector and supporting the public sector agencies in implementing the MfO initiative. These pilots, which were known as the Capability, Accountability and Performance (CAP) Pilot and the Pathfinder Project, were used as a means to support some departments that were willing to participate in the development of a more results-based approach to their policy advice and service delivery operations. The most extensive pilot project is the Pathfinder Project, which was implemented in 2001/02. There were 15 departments that took part in the Project, including some which had participated in the earlier CAP trials (Economics and Strategy Group, 2003). For other non-participating departments, support was provided later in the form of an intensive series of workshops and guidance materials under the MfO initiative (Webber, 2004).

The official announcement of the MfO initiative was made in December 2001, which was close to the announcement of the results of the RoC. MfO was intended to require all public sector agencies to adopt a more strategic and outcome-focused approach to management and reporting (The Steering Group for the Managing for Outcomes Roll-out 2004/05, 2003). In this new initiative, departments were required to produce an annual Statement of Intent, to replace the Department Forecast Report as the performance planning document, that would emphasise a 'strategic outcomes framework' (Webber, 2004). This document was supposed to function as a means for strategic discussions between agencies and their respective ministers and other stakeholders (Anderson & Dovey, 2003). The consequence of this focus was a gradual shift of reliance from the previous narrower output-based planning to a broader outcome-output-based planning (Webber, 2004), which would take into account the responsibility of executive agencies for the development of their long-term capability (Anderson & Dovey, 2003). This initiative essentially puts more burdens on chief executives to improve the quality of outcome information in public management processes.

With this new initiative, the public sector reform in New Zealand shifted from the question of 'what does government do?' to 'does government achieve anything?'

(Gray, 2005). Although chief executives were not accountable for achieving the outcomes (results), they were now accountable not only for delivering the outputs but also for their contribution to the achievement of outcomes. In their reporting documents, executive agencies were to show that their outputs were specified to achieve certain impacts (results), which were supposed to be agreed upon with their key stakeholders, and measures were described for achievement of these results. The SOIs and Annual Reports had to contain commentary and indicative measures on outcomes, strategies, risk and capabilities (Anderson & Dovey, 2003).

As this initiative was starting to be implemented, two issues with the existing output-based budgeting were also identified. First, the legislative requirement for departmental output based on 'Output Classes' had created some problems when ministers attempted to relate the output classes to the intended outcomes. There was some evidence that the expenditure categories contributed to an increasing lack of understanding and interest of ministers in the detailed spending of their departments (Economics and Strategy Group, 2003). As stated by Webber (2004), although there was a shift towards more management-related output categorisation in recent years to enable ministers to relate the outputs to the outcomes, the criteria were not consistently applied. Furthermore, only some departments provided a meaningful indication of the purpose and the intended impact of expenditure to Parliament or the public. Second, several significant areas of departmental administration were difficult to define in a meaningful and measurable way in terms of outputs and output performance, such as policy advice or ministerial servicing (Webber, 2004). As outputs could not easily be specified, the related outcomes were also more difficult to specify.

To date, two reviews have been conducted on the MfO initiative: one by the Treasury and the other by the Controller and Auditor-General (Economics and Strategy Group, 2003). The first review was conducted by Roger Beckett and Simon Gilmore in 2007, following a request from Finance Minister Michael Cullen. They conducted interviews with both parliamentary and non-parliamentary users on their perceptions of departments' accountability documents. They found that there was a concern among users that these documents were not as effective as they should be, in fulfilling their purpose of supporting decision-making in budgeting and assessing performance (Beckett & Gilmore, 2007). The second review assessed the quality of service performance information contained in the 2007/08 SOIs based on their developed criteria. Overall, the CAG was disappointed with the quality of information reported, especially in the presentation of logical links between the medium-term outcomes and

outputs planned and the specification and relevancy of performance measures and targets for the outputs and outcomes (Controller and Auditor General, 2008).

## **G. An Overview of Managing for Outcomes**

The MfO initiative gave more responsibility to chief executives. In addition to their responsibility to deliver outputs, chief executives were responsible for managing their department's activities toward achieving outcomes. Instead of waiting for the government (or ministers) to come up with the department's outcomes, chief executives were required to be pro-active in ensuring that the outcomes were set so that they were able to align policy advice and service delivery to pursue them. Chief executives had to show that the outcomes were established and measured and that they were altering their mix of outputs as circumstances dictated, to better pursue outcomes.

In the SSC's official documentation of MfO, the Economic and Strategy Group (2003) stated that the ability of chief executives to implement MfO initiatives was confirmed by the visibility of some key components. These key components of MfO were categorised in three aspects: planning, management and reporting. In the reporting aspects, changes in the requirements for issuing reporting documents (the Statement of Intent, Estimates, Output Plans and Annual Report) were minimal in the immediate future. There was no new document to be produced under this initiative. The only cosmetic change was the change in the name of a planning document title, from the Departmental Forecast Report to the Statement of Intent. However, over time, the content and quality of these documents were to reflect improvements from the adoption of the MfO initiative in the area of planning, management and performance. The documents were expected to show that chief executives were managing the delivery of outputs toward some specified outcomes.

There were to be significant changes in the planning aspect of public sector organisations. Three components that were expected to be visible in this area, as discussed by the Economic and Strategy Group (2003), were:

- **Strategic planning.** In strategic planning, the chief executive had to be able to relate the department to the sector(s) in which the department was operating by incorporating a well-defined set of government or sector outcomes. They had to be able to articulate what their department was trying to achieve in contributing to the broader objectives or outcomes. In their planning, they were essentially preparing an outcome framework with an effective intervention logic that linked their outputs to the government's outcomes. The most visible result from

improved strategic planning was to be in the form of the availability of a well-structured department's outcome framework in the corporate documents.

- **Stakeholders' consultation and coordination.** Chief executives had to be able to show that the stakeholders were involved in the outcome-focused strategic planning. The planning had to involve at least two tiers of management within the department. The relevant outside stakeholders were to be consulted to a certain degree, especially the responsible minister(s).
- **Improvement by trial and experience.** Since a learning process is necessary in the effort to find the most effective interventions to achieve certain outcomes, there was an expectation that over time, the quality of the outcome framework and the discussion around it would improve with trial and experience.

There were five visible key components in management by which the chief executives had to show that they were successfully implementing the MfO initiative. These components, as explained by the Economic and Strategy Group (2003), were:

- **Modified business plan and/or improved output plans.** There was an expectation that the business plan included the output plans, when departments were making a detailed operational plan of their new strategic framework. The progress was reflected in the form of alignment of the department's existing outputs to specific outcomes and/or modification of outputs to achieve the best combination to achieve the stated outcomes.
- **Implementation of performance monitoring.** Since outcomes were to be consistently monitored, there would be performance measures against outcomes, especially for the priority or vital ones. Departments had to develop a capability to monitor performance that would support the managerial decision-making over time.
- **Changes in the management system, organisational culture and/or long-term capability requirements.** The long-term reflection of managerial engagement around outcomes may be reflected in changes to these organisational features.
- **Improved coordination among government agencies and better communication with stakeholders.** Outcome-oriented management required chief executives to continuously consult and collaborate with relevant stakeholders, central agencies and other government departments. The evidence that they were going in the right direction was initially seen in the development of shared outcomes and later in the possibility of shared funding, managerial, or

service delivery responsibilities. If chief executives understood what other agencies were contributing to the broader goals, it would also show that coordination and cooperation among agencies had improved.

- **Improved internal coordination.** Internally there would be improvements in the coordination within a department. Staff would exhibit better understanding of their departmental objectives and their division's or business unit's contribution to the achievement of the departmental objectives.

## **H. The Current State of Performance Reporting in New Zealand**

In terms of information disclosure, two major changes have been implemented under the MfO initiative. First, executive agencies are required to provide outcome information in addition to information about their planned outputs in the SOI. Second, they are required to report the achievement in terms of output production and outcome achievement in the Annual Report.

The SOI is a planning document submitted by government agencies to Parliament that provides a description and explanation of the agencies' key operating intentions and performance expectations over the medium term of a 3-5 year period. The content requirements as stated in the PFA can be divided into two broad sets of information: a medium-term set of information for a minimum of three financial years and an annual set of information for the first financial year covered by the SOI. The medium-term set of information describes the agency's function, their targeted achievements (goals, impacts, outcomes and/or objectives), their strategy for achieving the targets, and their proposed metrics to measure the achievement. In terms of the strategy, agencies should describe the following information: the rationale (links) between their outputs and their selected and targeted outcomes, the risks (unintended outcomes) identified and managed, their past and future major evaluative activities to monitor their achievements, and any collaborative activities planned with other organisations (The Treasury and State Services Commission, 2007).

New unaudited information is also required to be recorded in the Annual Report relating to the achievement of the medium-term and yearly plans described in the related SOI. The PFA requires agencies to disclose their progress against the targets set out in the SOI. The reports should cover the achievement of agencies against their targeted milestones using the proposed metrics described in the SOI. The reports should also contain information about the execution of the strategy. Agencies will determine their own format for reporting this information since there is no particular format

prescribed. The Treasury cites the criteria used in the NZ Institute of Chartered Accountants Annual Report Awards as a reference for good reporting (The Treasury, 2006).

The current reporting regime for performance reporting in NZ consists of two sets of information: audited output information and unaudited outcome information. While the Statement of Service Performance could be used to report the output and outcome information, most agencies only reported output information in the Statement of Service Performance. The Statement has to be audited by the Auditor General in terms of the accuracy of information. Some documents can be referred to for the presentation, as the Treasury points out (see for example, The Treasury, 2006; The Treasury and State Services Commission, 2007), such as the Technical Practice Aid No. 9: Service Performance Reporting or the OECD's documentation entitled "Specifying Outputs in the Public Sector."

The outcome information is usually reported in the unaudited part of the Annual Report. For the reason of avoiding the audit requirement, most agencies report the outcome information in this part of the Report. Therefore, the emphasis of this study is the whole Annual Report, not just the Statement of Service Performance, since the former provides more complete picture on the presentation of output and outcome information. The format of outcome information presentation is presented usually in the form of narrative information with some graphical and diagrammatical information. Similar to output information, there is no accepted standard for the presentation of this information. More documents are available for reference in the reporting of this information. Those documents are:

- FRS-2: Presentation of Financial Reports, issued by New Zealand Institute of Chartered Accountants (NZICA)<sup>2</sup> in 1994 and amended in 1996. This document sets out, among other things, the mandatory requirements for service performance reporting. Until the introduction of International Financial Reporting Standards (IFRS) in the public sector in 2007, this standard was legally binding on the entities to which it applies since it had been approved by the Accounting Standards Review Board (ASRB) for the purpose of the Financial Reporting Act 1993. The corresponding standard under IFRS is NZ IAS 1 Presentation of Financial Statement (last revised in 2007 and amended in

---

<sup>2</sup> This body was first known as NZ Society of Accountants, the Institute of Chartered Accountants in New Zealand and now as New Zealand Institute of Chartered Accountants

2009). This standard has also been approved by the ASRB and is thus also legally binding on affected entities.

- TPA-9: Service Performance Reporting, issued by NZICA in 2002. This document contains comprehensive guidance for the specification, measurement and reporting of service performance. The guidance is developed based on best practice in service performance reporting.
- Reporting Public Sector Performance, 2nd Edition, issued by the Controller and Auditor General in 2002. This document contains the Auditor General's view on the proper presentation of reporting public sector performance.
- Managing for Outcomes: Guidance for Departments, issued by the Treasury in 2003. This guidance is intended to help government departments in implementing the MfO initiative. It contains suggestions on how outcome information should be reported in the corporate documents.
- Guidance and Requirements for Departments: Preparing the Statement of Intent, first issued by the Treasury in 2003, and amended yearly with the latest one being 2007. This document provides a detailed instruction on how to prepare the Statement of Intent, including the format of the presentation of outcome information.
- Public Finance Act 1989, as amended in 2004. This act sets out the requirements for public sector reporting. The act also lists the required information to be disclosed in the service performance disclosure. The 2004 amendment sets out further performance information related to outcomes, cost effectiveness and organisational health and capability.
- Preparing Annual Reports, first issued by the Treasury in 2005 and later amended in 2006. This document provides detailed guidance for departments to prepare their annual report, including the format of the presentation of performance information.
- Performance Information Measures and Standards in the SOI and Annual Report, issued by the State Services Commission in 2006. This document is intended to provide further and elaborate guidance for government departments in developing and reporting performance information in the corporate documents.

### **III. LITERATURE REVIEW**

#### **A. Theoretical Framework**

Public choice theory and agency theory are the most important economic theories underlying public sector reform. This research focuses on the influence of these theories on the current situation in the New Zealand public sector. The aim is to point out which components of the reforms that do not consider these theoretical perspectives. Therefore, this research uses these perspectives as the lenses for collecting and analysing data to support the findings. Public choice theory is used to identify and explain the conformity of the changes in the public sector system to those proposed in the theory, especially in the area relating to information produced in the legislative/political environment. Agency theory is the lens for identifying and explaining factors that led to the adoption and improvement of performance information reported by the New Zealand central government agencies. The focus of analysis using this theory is on the relationship between government agencies and various stakeholders in relation to the production of information. This theory is used to identify the presence (or the lack) of mechanisms used and their effectiveness in promoting more performance disclosure by agencies.

##### **1. Public Sector Problems in the Perspective of Public Choice Theory**

Public choice theory is a positivist approach to studying and understanding the political and bureaucratic process in the public sector. It is an application of economics to the political environment. This theory explains the incentive mechanisms and processes of the appropriation and distribution of goods through political means, rather than through purely economic means. The political means of public good production is unique, since the benefits and the costs can be concentrated or diffused. This theory is different from the previously dominant normative public sector approaches because of the difference in the assumptions of the participants in the political and bureaucratic spheres. The conventional view holds that public officials are pursuing public interest, in which they faithfully carry out the “will of the people”. In modelling the behaviour of individuals in the public sector, public choice theorists assume that people are guided predominantly by their own self-interest. Their motivations are no different whether they participate in the political process or in the market. Voters will support candidates



or ballot propositions that they believe will make them personally better off. Bureaucrats will seek to advance their careers. Politicians want to be elected or re-elected. In essence, public choice theory is the use of the rational actor model of economic theory to understand the realm of the public sector environment.

Ostrom & Ostrom (1971) list four assumptions of public choice theory:

- Individuals are self-interested, which means individuals have their own preferences that affect the decisions they make. These preferences may differ from one individual to another.
- Individuals are rational. It means they are able to rank all known alternatives available in a transitive manner.
- Individuals adopt maximising strategies, in which individuals make consistent choices from those alternatives that they think will give them the highest net benefit based on their preference.
- There are levels of information possessed by the individuals: certainty, risk, and uncertainty. These levels are based on the individual's knowledge of strategies, outcome for each strategy and his/her preference for each outcome.

Public choice theory focuses on explaining how decisions are made in the public sector and why the decisions taken may not be as optimal as those that are made in the market. The fundamental unit of analysis in this theory is the individual. Organisations or a group of people do not make choices, only individuals do. Therefore, the model of public sector decision-making should incorporate how the diverse and often conflicting preferences of these self-interested individuals are expressed and pulled together when making a collective decision.

This theory also describes why public choice processes are different from those in the private sector. The differences are in the incentives and constraints that direct the self-interest pursuit of individuals in these two sectors. In a private choice process, such as buying cars, individuals choose among the available alternatives based on their personal circumstances. They capture the benefits and bear the costs of their own choices. The purchase decision is voluntary and a deal will happen if both buyer and seller are better off after the deal. On the other hand, for decisions in a public setting, such as politicians proposing a programme to fund the fixing of leaky homes for New Zealand citizens, the benefits and costs will be shared. There is no guarantee that everyone is better off with this decision. People who have leaky homes will support this programme because they are likely to get most of the benefits. They will show greater support if the funding comes from taxes that are assessed uniformly on the community

as a whole. However, people who do not own leaky homes will vote rationally against the proposal, since the costs for them exceed the benefits. If they are the minority in the community, they will be forced into paying for it. The collective decision-making process will make the majority impose their preference on the minority, unless those who are harmed can relocate easily to another political jurisdiction or individuals can veto a proposal that would harm them.

Public choice theory depicts the behaviour of three actors in the political realm: voters, including interest groups; politicians; and bureaucrats. There are two hierarchies in which these actors interact with one another to determine the appropriation and distribution of goods. The first hierarchy is the political market in which voters, interest groups and politicians meet and are involved in an aggregate decision-making process to sort out government policies. Bureaucrats do not play an active role in this hierarchy. They may provide politicians (in the parliament) with information for the decision-making process, but they do not actively seek to change the outcomes of the process to satisfy their preference. In fact, in most cases, they are not allowed to actively participate in the political process. The second hierarchy is in the government management process, in which policies are implemented by bureaucrats. In this hierarchy, the main actors are politicians (ministers in the government under the Westminster system) and bureaucrats. To a certain extent, interest groups can indirectly influence the process through policy networks in which bureaucrats are exposed to them. In these networks, bureaucrats in some cases maintain a working relationship at arm's length on behalf of the ministers (B. Ryan, 2006). Public choice issues exist in these interactions, both in the political and the bureaucratic processes.

### **Public Choice Issues in the Political Process**

Government decision-making is affected by some internal and external factors. The internal factor is the self-interest of politicians and bureaucrats and the external factors are pressures from interest groups, lobbying and voting behaviour. Since politicians are usually elected to represent their geographical constituents or their special interest constituents, they have a tendency to pursue policies, such as supporting pork-barrel projects, that will benefit their constituents possibly at the expense of the general public.

Interest groups influence political decision-making if they can get disproportionately more benefits from the decision while the costs are shared proportionately or disproportionately more by other members of society. They have

much stronger incentives than taxpayers and other large groups do, because they are small in numbers but have access to a concentration of benefits. As Stigler (1971) argues, group power is grossly unbalanced in favour of small and concentrated interests. However, this is true only in cases where the benefits of regulation or political decisions are concentrated and the costs are diffused.

The behaviour of people in using government legislation and regulation to transfer wealth (rent) to themselves is called rent-seeking (Johnson, 1994), which is the public choice problem in the political market hierarchy. Special interest groups concentrate their efforts to get preferential legislation/policy at the expense of the public at large. Politicians produce and supply legislation/policies, which redistribute income and wealth. They allocate rents not earned in the market. The elected representatives will have a biased perspective, as they need to cater to the interest groups by pursuing policies that benefit this small number of people at the expense of the other, larger, group of people in society. In other words, government policy is captured by interest groups. If the policy involves the establishment of a new government institution, the result is captured bureaucracy, which is a bureau that is captured since its inception. The bureau is designed from the start to promote regulated interests. Despite these rent-seeking problems, the presence of interest groups does not produce entirely negative effects. They may benefit the political process by providing additional information that may improve the quality of the decision made and the equality between the parties to the decision (Johnson, 1994).

The rational behaviour of politicians, interest groups and voters has led to situations in which the results of political decision-making are outcomes that conflict with the preference of the general public. Many biased policies or pork-barrel projects are not the desire of the overall democracy. However, it is rational for politicians to support these projects or policies for several reasons. Their support for the demands of a specific interest group may give them certain benefits and (possibly) make them feel powerful and important. They may gain financial benefit in the future by becoming a lobbyist for these groups after they retire from politics. They may support the policies or projects because their local constituents will get the benefit, which in turn will increase the possibility of getting more votes or campaign contributions. Another key reason is that the politicians pay little or no cost to gain these benefits, because they spend public money. Interest groups also behave rationally in rent-seeking behaviour. The result of policies or projects in their favour is millions or billions of dollars of benefits to the interest group compared with relatively small investments made. For some business

interest groups, winning government favour means that they have a better chance against their competitors. Voters or the general public are also rational. They calculate that it costs them so much more if they have to defeat the government that produces biased policies, than to lose small benefits individually from these policies. Voters may not even be aware of what is happening because they choose not to become aware. This is called 'irrational ignorance'. Voters induce that the cost of becoming informed (with the potential of taking action) is greater than the benefits of being informed. This is one of the key differences between the political market and capital market.

Incorporating Williamson (1985)'s approach of transaction cost economics, Frant (1996) develops an analytical model of the public choice problem in the political market and offers a solution to the problem. With the perspective of transaction cost economics, there is a choice between conducting a transaction in the market versus moving the transaction inside a firm or organisation. The incentives that are provided by the market transactions are called high-powered incentives since the "efficiency gains from the transaction flow directly to the parties transacting" (Frant, 1996, p. 367). Incentives in hierarchies are low-powered ones. Frant (1996) extends the definition of high-powered incentive so it can be used in the public sector analysis. He argues that since politicians desire re-election, this is similar to the market participant's desire for monetary gain. If high-powered monetary incentives promote productive efficiency through innovation and cost reduction, such high-powered political incentives promote "allocative" efficiency by increasing the likelihood that people will value the public sector's products.

Frant (1996) further argues that while a powerful incentive can make people more responsive, it may also result in greater unintended consequences in some situations. When people have opportunities to behave dishonestly, high-powered incentives may induce greater dishonesty, such as in the case of the manufacturer and supplier in the private market. In the political realm, politicians may engage in exploitative behaviour if political opportunism exists. It is a function of the knowledge of voters about politicians and the ability for them to monitor politicians (Frant, 1996). Politicians tend to act unscrupulously when both factors are weak, to ensure re-election. For example, they may spend public money in an area that would not be selected if voters were fully informed.

The solution to the opportunism problem is to move the decisions into a hierarchy in which incentives are low-powered. In other words, if political opportunism strongly exists, the decisions of the government have to be depoliticised as far as

possible. The activities are put out of reach of the normal political process, so that the previously impossible transaction that provides mutual benefits can now take place. Legislators cannot easily undo a law or annul a contract from one legislative session to another. When the activities are assigned to an independent public bureau, a structure should be in place so that voters and politicians can commit themselves in advance to avoid certain types of undesirable behaviour.

### **Public Choice Problems or Agency Problems in Bureaucracy**

One drawback of low-powered incentives in the public sector is that they eliminate the allocative efficiency that is possible in the environment of high-powered incentives. It is harder to ensure that decisions or transactions in a bureaucracy reflect citizens' preferences, as the incentives are not designed to promote sensitivity to society's preferences. The issue is now about how to control the problems of politicians with bureaucracy. Frant (1996) considers that allocation decisions in the hierarchies may be motivated by professionalism norms and the decision-makers'/bureaucrats' perception of public interest. These motivations, coupled with an accountability mechanism that would make bureaucracy become responsive to the politicians' wishes, can be the right combination for promoting allocative efficiency of the government institutions' decisions.

Moe (1995) has a different perspective when explaining why the bureaucratic structure is not easily influenced by politicians, thus resulting in its unresponsiveness to the politicians' wishes. He starts his argument by considering that the right to exercise public authority is a sort of property right. Politicians use this right formally and interest groups use it informally by influencing politicians to make choices about government policy and structure. However, in a democratic society, no set of individuals can perpetually claim these rights. Some individuals may gain support today in order to be very powerful, but they are not sure that they can maintain their position of power in the future. There is uncertainty with political property rights.

This uncertainty has produced a profound effect on the behaviour of politicians and interest groups. They know what they gain today in terms of policies and structures can be overturned by actors with different interests tomorrow. Their hard-won achievements can be undermined or destroyed by others in the future. They have strong incentives to make sure that this situation will not easily happen. Ex ante, they will try to anticipate this by taking preventative actions. In the design of policies or structures, they will install a capacity for the policies or structures to survive and prosper in an

uncertain political future. The structures therefore involve the separation of politics from administration. In their desire to ensure that bureaucracy is unable to be influenced by their opponents, politicians have ensured their installed policies and structures are not easily controlled by them. In Moe (1995)'s proposition, the political unresponsiveness of government agencies has been designed from the inception of the agencies, because the politicians do not want their policies or structures to be easily overturned.

Once politicians set up structures or bureaucracy to implement their choice of policies, public choice problems occur in the bureaucratic process. There are three types of public choice problems. The first type is essentially the same as the public choice problem in the political market, which stems from the rent-seeking behaviour of interest groups. In fact, this problem originates in the political market and is entrenched within the bureaucracy. Some politicians want to ensure that bureaucrats are not subject to capture by their clients, who have different interests to the general public and in particular to their interest as elected representatives of the public in the current government. These clients are the interest groups that worked with the politicians in the previous government to capture the policies currently implemented by the government agencies. The politicians in the current government will find it difficult to change these policies and introduce others in line with their preferences. The bureaucracy is not responsive enough to the policy preferences of the governing politicians. The second type is called bureaucratic capture, which is the rent-seeking behaviour of the bureaucrats. Officials in government agencies have a tendency to pursue their own interests, which may not be in line with the interests of their political principals. The third type is capture by agency employees, which is essentially the agency problem of chief executives in employing staff to deliver services. Public sector agencies may be unresponsive or deliver services in an inefficient manner because employees are maximising their interests rather than their principal interests.

Moe (1997), extending Niskanen (1971)'s model of the relationship between politicians and bureaucrats, argues that the game played between the two focuses on asymmetric information and the legislature's authority to set the rules and assert control. In this game, bureaucrats try to maximise the budget slack allocated by the legislature using their superior information. In Niskanen's perspective, they have two pivotal advantages over politicians. Usually, they are the sole supplier of public services so they have a monopoly over information about the true cost of production. They also know the value of every level of output to the legislature. Politicians, on the other hand,

determine the rules of the game because bureaucrats are under their authority and politicians are the ones that set the structure of the bargaining process (Moe, 1995).

One course that politicians can take to reduce the bureaucrats' tendency to engage in rent-seeking behaviour is to develop and implement intensive monitoring. However, according to McCubbin & Schwartz (1984), they have little incentive to engage in this kind of monitoring. Instead, they prefer to use a "fire alarm" strategy to monitor bureaucrats. This is because it costs them more to engage in "police patrol" oversight without any meaningful additional benefits that they can gain. Their incentive is to be re-elected. They just need to satisfy their constituency groups to win their votes by simply responding to any fire alarms these groups "set off," and identify what it is when something goes wrong. McCubbin & Schwartz (1984) further argue that responding to their voices makes sense for two reasons. First, it enables politicians to produce tight control by combining their response with their legislative weapons. These weapons are powerful enough to shape the bureaucracy. Because of this huge consequence, bureaucrats will tend to anticipate it and comply from the outset. Second, the constituent groups feel that politicians do address their concerns and therefore they are more likely to support the politicians in elections. These groups are most likely the customers of the bureaucracy. They cannot discipline the public sector agencies directly, since there is no exit strategy available if they are not satisfied with the services provided. The only mechanism they have is voicing their concerns to the politicians, asking them to act in disciplining the agencies in return for their votes.

The weakness of disciplining bureaucrats through the consumers' voice is that politicians filter demands through legislation. Public sector agencies may not be able to judge the consumers' wishes because of this filtering. They may try to open their own information channel to consumers, exposing them to capture or forcing them to seek capture. In addition, these agencies cannot be disciplined properly to meet the demands of the consumers, as they may over- or under-estimate the demands. Furthermore, when politicians try to resolve the voters' specific demands/complaints, it induces public sector agencies to do political micro-management.

## **2. Agency Theory and the Public Sector Setting**

Agency theory is a theory that discusses the ubiquitous relationship between two parties in which one party (the agent) performs work delegated by another (the principal). An agent is employed to undertake some activity on behalf of a principal.

Agency theory provides a theoretical basis and general framework for the behaviour of people with different goals in a principal-agent relationship.

Lambert (2001) describes the role of both parties in the relationship. The principal delegates tasks, provides capital, takes risks and develops incentives. The agent has to complete the tasks, make decisions on behalf of the principal and bear a secondary type of risk. The principal will seek to control and shape the behaviour of the agent so he/she will act in a manner consistent with the principal's preferences. One of the tools used to accomplish this goal is entering into a (compensation) contractual arrangement where the principal and the agent must agree on what the agent is paid, which includes how much compensation exists for the agent and what conditions apply to the compensation.

This standard agency theory has some important assumptions. Liu and Mills (2007) summarise these major assumptions as follows:

- The parties involved are self-interested. Principals and agents will seek to maximise their returns. Their utility functions are mutually independent, so the interests of the principal and the agent are not always aligned.
- There is goal conflict between the parties involved. The utility function of the agent has two major conflicts: income that provides utility and effort that provides disutility. Agents will always try to maximise income and minimise efforts. If an agent minimises his/her efforts, it creates disutility to his/her principal, who seeks to have the agent do his/her best effort. Thus, the agent may not act in the interest of the principal.
- There is significant information asymmetry. Principals have perfect information on what actions should be done by agents but there is information asymmetry on the actual actions taken by agents. Only agents have the information of their actual actions, but some actions are observable to the principal.

Eisenhardt (1989) adds some important assumptions on the nature of the parties involved and the contract. The actions of the principal and agent are endogenously derived and based on well-specified preferences and beliefs. They also have unlimited computational ability and they can anticipate and assess the probability of all possible future contingencies. It is assumed that the contract is complete. The contract specifies actions for each verifiable situation. The courts are perfect enforcers of the contract.

The environment and the characteristics of the principal-agent relationship depicted in the above assumptions lead to the basis of an agency problem. The inherent problem in this relationship is that the agent may not act in the best interest of the



principal, which stems from the self-interested nature of the parties involved and the presence of goal conflict. This problem intensifies under the conditions of information asymmetry and the low efficacy of the agent, which leads to two specific types of agency problems: moral hazard and adverse selection. Moral hazard is an agency problem on the part of the agent's actions, in which there is a lack of agent's effort. When the interests of the principal and the agent are not aligned, the self-interested agent may take advantage of the principal by pursuing his/her own interest at the expense of the principal. It arises because of the presence of an information asymmetry in which the agent's actions are either hidden from the principal or are costly to observe. On the other hand, adverse selection is a problem relating to the principal's action in appointing the agent. This problem also arises because of information asymmetry. As a result, the principal cannot assess the competency of the agent before entering into a contract. The agent may be tempted to misrepresent his/her competency or ability in order to get the contract with more favourable terms. Therefore, the principal may unwittingly enter into a contract that otherwise he/she would not do if he/she could get perfect information on the agent (Eisenhardt, 1989; Waterman & Meier, 1998).

The principal develops strategies to help mitigate agency problems. Consequently, the principal has to incur some costs, which are called agency costs, and can be classified into two types: monitoring costs and bonding costs. Monitoring costs relate to the efforts to reduce information asymmetry, such as mandatory audit, development of governance structure, installation of information systems and other oversight processes. The main monitoring method used is requiring the agent to report his/her performance or behaviour and verifying the information by employing an auditor. Bonding costs relate to the initiatives to reduce goal conflict between the principal and the agent. The main method in these efforts is the use of an incentive scheme or reward structure in combination with performance assessment. The incentives may be explicit, such as receiving a bonus or getting more resources, or may be implicit, such as a consideration for employment (Eisenhardt, 1989; Waterman & Meier, 1998).

The agency problems become more severe in certain conditions which are more prominent in the public sector. Therefore, in analysing public sector cases through the lens of agency theory, the simple model of principal-agent relationship should be expanded to consider the presence of these conditions. The conditions, which will be explained later, are: less verifiable or measurable outcome/performance, ex-ante causal ambiguity, the presence of multiple principals, severe bounded-rationality on the part of

principals, a high degree of risk-averse behaviour on the part of agents, and the level of goal conflict in the relationship.

The verifiability or measurability of outcome will influence the monitoring mechanisms used by the principal. If outcome dimensions are not measurable, information on performance by default is useless in a single-period agency relationship. Principals will tend to use behaviour-based monitoring assuming they know the types of agent behaviour that will result in better performance. In a multi-period relationship, principals and agents may learn to develop some verifiable but imperfect aggregate measures of outcome. This information will be used both for incentive purposes and for risk-sharing purposes (Banker, Datar, & Maindiratta, 1989). Implicit or relational contracts also play an important role in this situation if the value of the future relationship is large enough not to induce either party to renege (Baker, Gibbons & Murphy, 2002). In this kind of contract, subjective (unverifiable) measures of performance are used, together with available objective measures, to disentangle the effects of an agent's actions from those of luck and to eliminate distortions induced by objective measures of performance. On the other hand, when the dimensions of performance are ex-post verifiable then the monitoring system should be able to capture the performance information. More performance information will be demanded in a situation in which it is more verifiable. Therefore, the greater the measurability of outcome the more performance information will be used in a principal-agent relationship.

The second condition is the existence of ex ante causal ambiguity, in which the principal does not know what specific agent actions add value to the relationship. Jacobides & Croson (2001) argue that the principal is in a difficult position, since he/she has to commit to a scheme to assess the usefulness of the agent's effort before learning about it. Since the principal does not know what the right effort is, the monitoring system is doomed to fail, as the information generated lacks relevance. In a multi-period relationship, the principal may ultimately gain adequate understanding of the relationship between outcome, efforts and chance factors. However, if the relationship is very complex, it will take a long time to have an adequate understanding that enables the principal to develop a management scheme for the agent's compensation.

Enforcing monitoring in situations in which ex ante knowledge of critical success factors is very limited will impair the organisation's survival ability because the learning environment is not supported. Principals and agents need to learn these factors over time. If agents are forced to be accountable or compensated on the basis of specific

and contractable dimensions, they have less incentive to advance a beneficial but costly learning activity (Jacobides & Croson, 2001). Agents will never be induced to experiment with new ways of doing things that will be more effective in achieving outcomes. Principals will never gain an adequate understanding of outcomes, efforts and chance factors. The organisation is doomed to fail because it will never be able to fulfil its objectives.

The next condition is the presence of multiple principals, which is a typical characteristic of the principal-agent relationship in the public sector. A public agency deals with two or more principals in a complicated way to produce public goods and services. It may have to deal with more than one minister, address the concerns of parliamentary members, serve the public and follow the directions of a coordinating agency. In dealing with the agent, some principals may focus on delegating tasks to achieve certain outcomes or to produce specific products or services. Other principals may have a bigger role in influencing resource allocation to the agency in the budgeting process, while others may be involved in specific activities such as the monitoring process and incentive contract management.

The presence of multiple principals will likely lead to a situation where there are multiple and/or conflicting goals. As different principals may have different goals, multiple conflicting goals are inevitable. These conflicting goals make the principal-agent relationship very complex. One or more principals will be dissatisfied with the relationship no matter how well principals design the monitoring systems and the incentive structure (Waterman & Meier, 1998). Some compromise may be developed, resulting in multiple and sometimes ambiguous goals or “noise”, as competing principals will rarely agree on the goals of the agency or on accounts of the agency’s performance. The agent will have limited knowledge of the preferences of the principals since there is no precise knowledge of them and uncertainty about the attention of these (political) principals (Waterman & Meier, 1998).

The design of incentive schemes is very complex in this situation. It is expected that there will be multiple sources of guidance, backed with different kinds of incentive mechanisms (Eisner, Worsham & Ringquist, 1996). These mechanisms compete with each other (Martimort, 1996). Dixit (1997), as quoted by Verbeeten (2008), argues that delivering the incentive scheme is very complex in this situation, as each principal will have positive coefficients only on the performance dimensions of his/her interest, and negative coefficients on other dimensions. Therefore, the incentives are weak, as the

aggregate marginal incentive coefficient for each outcome decreases with the increasing number of principals (Verbeeten, 2008).

The best strategy for these principals in monitoring the agent is using what McCubbins & Schwartz (1984) term as “fire-alarm oversight”, instead of continuous monitoring. This oversight relies on the voices of their constituents or the public in informing them that the agent is not acting in the constituent’s or public’s best interest. The incentive structure faced by the (political) principal for monitoring activities is the cause for adopting this strategy, as explained before in the discussion on public choice theory. If the agent continues to shirk responsibilities when the level of monitoring is low, the principals individually are unlikely to directly bear any costs of the shirking. The bulk of the cost will be passed along to the public as the ultimate principal of these (political) principals. It will only cost them if their constituents or public ask them to act to reduce the shirking in exchange for the constituents’ support or votes. Therefore, as long as the public is not aware (i.e. a scandal has not emerged), the political principals would not pay a direct cost for this lax oversight (Waterman & Meier, 1998).

The fourth condition is the problem of bounded rationality. This problem exists when principals (and agents) have limited computational ability and they cannot anticipate and assess the probability of all possible future contingencies. One of a few situations in which a bounded rationality assumption exists is when principals can specify their objectives but they may not be able to communicate them perfectly to agents. As they resist the specification of objectives, informal dialogue using examples and generalisation is the preferable method of the principals’ communication. They want their agents to be proactive in responding to complex or unforeseen circumstances. The agent’s initiative and judgment should be executed in the principal’s accepted way and in conformity with their values. Therefore, the agent’s ability to meet the principal’s objectives is a function of their honesty, competence and empathy with the principal (Hendry, 2002).

Another aspect of the principal’s bounded rationality is restriction on communications. The restrictions will make the principals acquire only limited knowledge of agents’ actions even though the information is fully reported by agents. One such restriction is technical expertise. Principals may not have the technical expertise necessary to understand the agents’ information provided (Lambert, 2001), because they lack the necessary expertise in complex areas or because they have bias in the information flow due to the “fire alarm mechanism” (Eisner et al., 1996). In this situation, the principal and agent relationship operates under greater uncertainty than in

the “full information” case, in which the principal is able to know for certain whether the actual outcome is due to the agent’s performance or due to factors beyond the agent’s control. This communication restriction also limits the ability of agents to truthfully communicate their information (Lambert, 2001). Even though the agent may be able to comply with the requirement to report an extremely rich information set, the usefulness of the information is limited by the ability of the principal to understand it.

Several situations stem from the condition of bounded rationality. First, this condition will make principals resisting the specification of goals or objectives, so that agents are unable to have an adequate understanding of those goals and objectives. This condition is similar to the one in which outcomes are not easily measurable. When the principal’s objectives are not easily understood, it is difficult to come up with aggregate measures of outcomes that show the achievement of the objectives. Less performance information is demanded in this environment. Even if there is no goal conflict, agents cannot be induced to reveal more, as they do not have the ability to understand all dimensions of the performance that show the fulfilment of the objectives. Second, since principals cannot understand the agent’s information fully in order to capture the full dimensionality of efforts or outcomes, they will not rely on periodic formal reporting in communicating outcomes or monitoring an agent’s effort. Instead, they prefer continuous informal dialogue and place more value on the agent’s initiative and judgment in signalling his empathy with them.

The fifth aspect is the characteristic nature of risk-averse agents. While parties involved in the principal-agent relationship are assumed to be risk-averse in a typical agency model, the risk-averse characteristic is stronger in the case of agents employed in the public sector setting. The risk characteristics of the agent will be important in an agency relationship when there is a stochastic term attached to the agent’s output. The agent faces risk in terms of uncertainty of outcomes, as outcomes are only partly a function of effort. Risk-averse agents will avoid any relationship where the incentives are based on outcome. Even if the agent’s performance is to be measured by the results of his/her effort, the agent will prefer measures that are controllable by them, whether the controllability is related to his/her ability to achieve the benchmark or their ability to manipulate the measures.

Another consequence of an agency relationship with a risk-averse agent is in the agent’s selection of his/her behaviour or efforts. Wright et al. (2001) argue that risk-averse agents prefer risk-reducing strategies because they can lower the disutility associated with their work and also reduce the prospect of losing their jobs. As

outcome-based incentive strategies are not preferable, principals have to rely on monitoring strategies. Only if tasks are specified clearly in advance is the principal able to induce efforts by using these strategies. In multi-tasking situations, risk-averse agents will be more likely to reduce their efforts on some non-visible tasks if they already exert high efforts on other visible tasks, even though the former tasks will be likely to produce better outcomes for the principal.

The last situation is the level of goal conflict in the principal-agent relationship. In a typical principal-agent model, there is divergence between the principal's and the agent's interests. Many principal-agent relationships in the public sector exhibit goal-conflict characteristics. Public sector agents have ever-changing relationships with different groups of principals, as politicians and political coalitions change. It is highly likely that they will have political principals who have divergent interests. Even if they deal with the same group of (political) principals who share convergent interests, in order to maximise their chances of getting re-elected, these politicians may seek to alter established policy in a new direction, which may result in the divergence of interests.

In other situations, the interests of both parties in the agency relationship may align with each other. This goal convergence or consensus might occur in the culture of collectivism that is prevalent in the public sector. In this culture, collectivist behaviours have higher utility than individualistic self-serving behaviours. Agents will be more likely to act in the interest of principals in this environment. Wright et al. (2001), using McClelland (1960)'s theory of needs, argue that some agents may enjoy performing responsibly because of their need for love, respect and self-actualisation associated with their employment. In this environment, the assumption of goal conflict may not hold. Verbeeten (2008) claims that the fact that more risk-averse employees tend to work for public sector organisations may imply that there is a match between them and these organisations. They may be motivated to work in the public sector by some intrinsic motivation, such as the idealistic or ethical purpose served by the organisations. Therefore, agents in these organisations may get utility from some aspects of the task itself, resulting in a more supportive attitude towards their principals' interests.

Waterman & Meier (1998) argue that there are two consequences seen in the agency relationship as a result of goal consensus, depending on the possession of information. Agents will play a passive role if neither principals nor agents possess information. They will support the policy proposals of the principals and protect and implement them without questioning them during the policy adoptions. Both may put additional effort into providing more information if it will help them clarify the

principals' interests. If information asymmetry exists, agents will voluntarily provide more information to clarify the goals. Once the goals are clear, tasks are delegated to agencies and they are simply left alone, and the principals will monitor them using a "fire-alarm strategy". Thus, in the situation of goal consensus, agents may be more easily induced to reveal more information.

### **The Role of Reputation in the Agency Relationship**

Reputation has been argued as being one important solution to the agency problem in which the scope of contracts is limited due to the circumstances of the agency relationship. Reputation has a potential for disciplining the behaviour of both parties in a multi-period setting. For the sake of reputation, agents as well as principals may both uphold their contract agreement, and fulfil any promises made under the implicit contracts they have agreed on during the course of the relationship. Thus, understanding how reputation works in the multi-period agency relationship is important in order to analyse public sector issues using agency theory. The multi-period dimension and the difficulty of providing explicit incentives in the public sector setting have made the role of implicit incentives, such as reputation, become more prominent in addressing the agency problem. Consequently, the study of reputation in the agency relationship may provide an interesting basis for the analytical analysis of agency relationships in the public sector.

Some researchers argue that the agent's reputation can be seen as his/her asset (Charmichael, 1989; Parker, 2005). The cost of this asset for an agent is the amount of shirking forgone in the past and when current efforts are made at more than a minimal level. The agent's reputation will potentially produce benefits in the future in terms of continuing and/or increasing streams of future income. The stream of future income is expected because of job security (i.e. the agent is not fired as he/she continues to have a good reputation), re-employment (where the contract is renewed once the old term ends), or the possibility of better job offers or positions. This reputation will remain valuable as long as the present value of the future earnings stream is greater than the benefits obtained through one-time cheating. Thus, as Parker (2005) argues, there is a quasi-rent for honesty, which is a price premium for preventing cheating in the form of a higher future income stream. Reputation expands the career horizon of an agent and reduces the attractiveness of opportunistic behaviour.

Other researchers view reputation as a mechanism for controlling self-interested behaviour based on social pressures. Society pressures an individual to conform to

certain norms and individuals obtain utility from developing a reputation that is consistent with those norms. Dees (1992), quoted by Stevens (2002), observes that social norms such as honesty, trustworthiness, fairness, justice, a sense of public duty, respect for the autonomy of others and avoidance of gratuitous harm, are the behaviours expected from society (the market) of a reputable agent. Therefore, if the agent wants to gain a good reputation, he/she must conform to the types of behaviour dictated by society's norms and consequently may get the benefits from it, such as entering into a less costly contract with a principal or obtaining more preferential treatment from the principal in the case of contract renewal.

An important idea arising from this career concern model is that even in circumstances where explicit compensation based on performance is not offered, the agent still has an incentive to provide more effort because outside options, in the form of future job offers from other principals, matter. Agents will try to convince the labour market that they possess a high level of talent by raising the market's perception of their ability. They do this by delivering a higher level of performance. This translates into future job opportunities with the same principal or other potential principals for potentially higher future wages. The agent will exert more unobservable efforts to increase the outputs and influence the market's belief. Nonetheless, in equilibrium the market will anticipate these actions and therefore draw a correct inference regarding the agents' ability, based on the observed outputs (Gibbons & Murphy, 1992).

### **Agency Relationship with Superior Private Information**

Cases in which agents have superior private information are easily found in many principal-agent relationships. Successful managers are said to have a comparative advantage in information acquisition as one of their prime skills. In the business setting, for example, a manager may hold information on productivity level, production costs or the market. Similarly, in the public sector setting, the manager may hold superior information on the functioning of bureaucratic production, the productivity of the capital employed and the characteristic of the public segment that his/her agency is currently serving. This superior information is gained because of the proximity of the manager to the production process and the accumulation of their experience in delivering public services. A typical principal in the public sector does not have experience of working or expertise in the area of public service production. Therefore, principals in the public sector have to deal with this situation of severe information asymmetry.



One of the consequences of the presence of an agent's private information is that if the information is obtained before the agent makes a decision, this additional information may improve economic welfare because of better selection of production choices by the agent that will benefit the principal, as shown in the hurdle model explained by Christensen & Feltham (2005). In this model, to have a higher probability of getting a good outcome, the agent has to overcome challenges (hurdles). If pre-decision information about the challenges is not available to the agent, he or she must provide the same effort for all challenges. With pre-decision information, the agent can make better decisions. If the challenge is too high, he or she will provide no effort. On the other hand, if the challenge is achievable, the agent will provide just enough effort to overcome the challenge and achieve a good outcome. The key issue for the principal is how to induce the agent to communicate the challenges truthfully.

The principal may be able to induce the agent by employing strategies based on the Revelation Principle in which there is a pre-commitment on the part of the principal to "under-utilising" the information. This principle states that "for any optimal contract (consisting of compensation, action and message) based on communication by the agent, there is another equivalent contract that (weakly) induces full and truthful disclosure of the agent's private information" (Christensen & Feltham, 2005, p. 262). In other words, according to this principle, any proposed mechanism involving non-truthful reporting by the agent can be duplicated or beaten in terms of expected utilities by an equilibrium mechanism that induces truthful reporting (Lambert, 2001). This principle does not say that truthful reporting is free. The "cost" of the principal's inducing the agent to tell the truth is that he/she has to use the information to a lesser extent than if the truthful reporting was not mandated or if there was a verifiable report with information. In some extreme cases, the principal has to ignore the agent's reported information in the compensation scheme. The revelation principle basically states that the cost of motivating a truthful reporting strategy is no greater than the cost of motivating a non-truthful one.

Lambert (2001) describes several types of models that incorporate conditions in which the revelation principle cannot hold, in his discussion on earnings management practices (non-truthful reporting). The first type is if the agent cannot report truthfully due to exogenous restrictions on the agent's ability to communicate his/her information. The second one is where the pre-commitment made on how to use the agent's report cannot hold.

According to Lambert (2001), one form of communication restriction is when the agent observes a rich information set but he/she is unable to communicate it fully and/or the principal does not have the technical expertise to understand it. In this situation, it will be very difficult and costly for the agent to communicate the full dimensionality of his/her information. In addition, the principal will also find it difficult to understand many dimensions of the information set reported. Here, by definition, the revelation principle fails since truthful reporting is impossible to achieve. This situation is similar to the bounded-rationality problem previously discussed. Thus, principals cannot use strategies based on this principle to induce the agent to disclose more information.

In the second model, Lambert (2001) shows that the revelation principle does not hold if the pre-commitment assumption is relaxed. When this pre-commitment cannot be enforced, the agent will believe that the information reported will be used against him/her. Then, it may be impossible to motivate the agent to report truthfully. Lambert (2001) argues that the pre-commitment assumption can be relaxed in two ways. In the first setting, there is the presence of another party (other than the principal and the agent) who cannot pre-commit on the use of reported information. This third party can be an auditor hired by the principal, another employee of the organisation, a competitor, or the labour market. In the second setting, the principal cannot pre-commit to the subsequent periods of a multi-period setting.

### **3. Public Sector Reform in the Perspective of Public Choice Theory and Agency Theory**

In public choice theory, the self-interest of politicians and bureaucrats is emphasised. The public sector will grow as government officials try to expand their budgets to get more budget slack, when politicians use public money to the benefit of their interest groups at the expense of other groups, and when the regulations and subsidies developed reduce economic growth (Boston, Martin, Pallot, & Walsh, 1996). The production of public goods is excessive because of budget maximisation behaviour, since there is no automatic disciplining mechanism available, such as market forces (Niskanen, 1971). Since the reward system in the public sector does not promote effective performance (Chapman, 1979) and environmental uncertainty makes contracts incomplete, principals are unable to monitor agents' behaviour (Dixon, Kouzmin, & Korac-Kakabadse, 1998) and agents act as rent-seekers. They seek privileges in order to

partake in the monopoly rent that they provide and reduce the efficiency of the economic system (Tullock & Eller, 1994).

Public choice theory offers a superior way of organising the production of public goods. The system should be based on some wider view of welfare in which everybody, including the producers and consumers, gets equal treatment. The system is based on minimal government intervention and closer attention to how individuals interact with each other. The right policies, in the public choice perspective, are the ones that minimise the government's role, control the discretionary power of politicians, reduce public monopolies, limit the functions of government agencies and maximise liberty (Johnson, 1994). Limiting the politicians' authority, such as preventing them from running budget deficits or imposing taxes beyond a certain level, may mitigate public choice problems in the political market. In the bureaucracy, government agencies can be constrained from implementing bureaucratic capture by limiting their authority to certain functions only. The role of policy advice provision should be separated from the role of implementing policy so that the advice given will not be biased towards the government agencies that would result in their bureaucratic capture. Competition should be developed in the provision of policy advice and public service, in order to provide consumers with more power to discipline government agencies in the form of an exit strategy, in addition to the impact their voices might make.

Public sector reform should also be directed at increasing the accountability of various actors in the public sector to their principals. The relationships between the public (via their representatives in the Parliament), ministers and executive agencies should be clearly defined. Executive agencies should be given clear objective functions, clear identified tasks and the right incentives to perform these tasks and to pursue these objectives.

The central assumption is that if executive agencies are given the right incentives, they can deliver socially desirable outputs or outcomes without the need for either direct political control or any administrative control. The public sector environment is then structured to improve incentive systems, contracting mechanisms and monitoring systems. Incentive systems should be improved by aligning them with proper performance measurement systems that capture the intended performance (behaviour) of agents. Contracting mechanisms are enhanced by providing ex-ante performance information on the expectations being sought from the agents. Monitoring mechanisms are advanced by providing more comprehensive ex-post information on agents' performance.

The optimality of the political market, just as in other markets, is based on the degree of access to information. The political market will be optimal if everyone has equal access to information. From the perspective of public interest, the quality of decisions produced by the policy decision process in the public sector will improve if there is a greater supply of relevant information to the actors involved. When transaction costs vary for different groups in the public sector realm then it is likely that access to information is not equal. Public sector reform should be directed at ensuring that more information is available to all relevant parties and that access to such information is equal for everybody.

The role of performance information is, therefore, critical to the reforms, as these improvements are related to the prescription of a better performance management framework. The reforms should be able to produce proper performance measures over time, as principals and agents are able to formulate better measures given their experience. These measures should be comprehensive and well specified. The annual report should contain the necessary performance information for monitoring purposes. If a reform is to be seen as successful in its implementation, the reform should exhibit these improvements over time. The quality of the performance information increases as the information becomes more comprehensive and well specified. The message from public choice theory is that if the reforms can restructure the public sector environment so as to have clear relationships among parties, proper incentive systems, more comprehensive contracting and monitoring, then public choice problems would be minimised.

The need for information and the difficulties associated with asymmetric information is a central issue in agency theory. This problem is inherent in the public sector setting, stemming from the delegation of legislative authority to executive agencies. Opportunities and incentives are created for agents to deviate from their principal's preference, as they usually have expertise that their political principals lack. Laws and principal directives (contracts) are not sufficiently complete and therefore they are open to an agent's interpretation. To align the agent's interest to their principal's preference, various mechanisms may be used, focusing on regulating the production of information, the use of incentives and improvement in monitoring. Some examples of these mechanisms are output-based commissions, profit-sharing, performance measurement and bonding agreements.

Agency theory has influenced public sector reform in the way that principals, their agents and the relationship between them are defined, including the framework for

incentives and monitoring. In the New Zealand setting, ministers are seen as the principal for executive agencies and the relationships are divided into two types: ownership and purchase interest. Financial performance measures are used for the incentive framework in the ownership relationship. Non-financial performance measures are the tools for aligning the agent's interest in the purchase relationship. Contracts, in the form of planning documents, and the subsequent performance information in the annual reports, are used for better monitoring of the behaviour of the agents.

Contracts, as the main tool for formalising principal and agent relationships, have the purpose of allocating risks, responsibilities and rewards precisely between the parties. In a principal-agency relationship, a contract has some important roles: aligning the interest of agent and principal, inducing agents to reduce information asymmetry using incentives, and providing motivations for agents to perform the desired tasks. In neo-classical theory, there are some implications for contracts that should be noted. Parties to the contract are assumed to be able to hide information and to cheat if opportunities exist. Therefore, contracts should be complete enough to cover all eventualities. Parties should easily be able to invoke the sanctions or rewards outlined in the contract. And lastly, the impact of the sanctions or rewards should be effective (Greve, 2000).

The insights gained from the perspective of agency theory are similar to the results of the previous analysis under public choice theory. Reform is said to be progressing if mechanisms exist to enable principals and agents to develop better contracts and the related monitoring mechanisms. Better specification in the contracts will result in better quality of performance information provided in the reporting documents. This information in turn will make the next period contracts contain more comprehensive performance information. The typical benefits associated with this improvement would be enhanced control and accountability, and reduced uncertainty and information asymmetry.

In terms of the reporting improvement, agency theory provides an insight into the pattern of performance improvement over time in the relationship between the agency and its principal. There is a problem of a *ratchet effect* in a multi-period contract setting. This effect would be exacerbated if there were no comparative performance measures. As the principal tends to consider the agent's past superior performance as a guide to evaluating his/her current performance, the agent is induced to lower his/her performance in earlier years to avoid being assessed against a higher standard in the

future (Milgrom & Roberts, 1992). The reporting improvement is influenced by the agent's contract period. The improvement would be higher in the latter periods of the agent's contract.

#### **4. Performance, Performance Information and the Users of Performance Information**

It can be argued that the performance reporting is a necessary element in the delivery of superior performance within the public sector. Therefore, performance information is always addressed in the public sector reform, even though the focus of the reform is the increase in the performance of public sector agencies. To promote superior performance in the public sector, high-quality performance information is needed. This information plays two important roles in the process of promoting superior performance: to secure the control of public sector principals on public managers and to provide incentives for the managers to perform better.

On the side of the principal, one of the arguments for the publication of high-quality performance information is the desire to secure control of the public sector organizations (M. Smith & Taffler, 1995). Control can be achieved if the public sector principals (the ministers, members of parliament or taxpayers) can assess the performance of the public sector agencies. The logic is when public and their political agents learn that the performance of these agencies could be improved, they will exert pressure on management to deliver better performance (P. Smith, 1990). In order to assess the performance, the principals should be able to identify objectives pursued by the agencies and the achievement of these objectives by the agencies. Performance information disclosure facilitates this process by communicating relevant information to the principals (Lee, 2008). The disclosure of this high-quality performance information will help the public sector principals to discipline the public sector managers by requiring them to continuously improve their performance if they do not want to be replaced. Performance information is said to be high quality if it can show the achievement of the objectives of the public sector agencies.

The publication of high-quality performance information also pressures public sector principals to manage their portfolio of public agencies more intensively. James & John (2006) argue that incumbent elected agents are often held responsible by voters for the performance of public sector agencies. When high-quality performance information is published, voters and service users may be able assess how well their public sector agents controlling the public agencies. The visibility of these agencies' performances

will result in the desire of the elected officials in charge of the government to exert pressures on the public sector managers to promote superior performance.

On the side of the agent, the publication of high-quality performance information can also work as an incentive mechanism for the public managers to promote superior performance. This is essentially the role of reputation as incentive mechanism in the principal-agent relationship that is already discussed. As Mayston (1985) argues, when relevant performance expectations are set, the subsequent performance information publication may produce ‘embarrassment effect’ on behalf of the public sector managers from failing to meet the performance target. This may then stimulate the managers to search and experiment for future improved performance. If the set of performance information published represents a complete set of information on the agency’s performance, the publication of high-quality performance information may potentially promote superior performance within the public sector organizations.

Performance information provision is an important mechanism in the process of mitigating the agency (as well as public choice) problems. The desire to increase the accountability of various actors in the public sector setting has led to a greater demand for performance information. The production of this information needs to be justified in a sense that its benefits should outweigh its costs, especially on mandated disclosure in which the users may not bear the production costs of the information. The performance information may be used as a part of control mechanism or incentive scheme to ensure that the executive agents deliver improved performance over time. However, the type of the users of performance information and their characteristics may influence the usefulness of performance information disclosure.

There is an established literature that discusses the users of public sector reporting. Most of earlier works in this area, such as research done by Anthony (1978), Jones et al. (1985), and the Government Accounting Standards Board (1987), were directed to identify the possible users in the public sector financial reporting. The later works were built on these earlier works to provide the similar list of users for non-financial performance reporting.

Mayston (1985) provides a comprehensive list of users of performance information disclosures in a form of four groups of relevant decision-makers that use performance indicators. The first group includes voters, tax-payers and consumers of goods and services produced by public sector agencies. The second group consists of elected officials, such as MPs and select committees, who are the representatives of the first group. The oversight officials, such as Audit Office and Ombudsmen, could be

included in this group since they provide oversight services to the Parliament. The third group includes the elected officials in charge of the governments (the ministers) and their support people (advisers). And, the last group consists of internal users, which are managers of the public service agencies.

There are several other users that are not identified in the list. Media and educational and research institution may get the benefits from the performance information disclosure since they can act as the information analysts for the other users. Other public sector agencies (within a jurisdiction or in different jurisdictions) are also identified as a significant user of performance information disclosure (Mack & Ryan, 2006).

The users' interest in the performance information is not homogenous. In general, the more focus the interest, the more intensive the demand for performance information disclosure and hence the more useful is the disclosure. The focus of the interest relates to the levels in the chain of accountability.

At the lowest level, in which the ministers in charge of the public sector agencies are the principal, information is expected to be used extensively as the ministers will need it to help managing their portfolio of public sector agencies and the consideration of funding support for programs in the portfolio (Pollitt, 2006). While the published performance information may be useful for the ministers' decision-making activities, they will be more likely to demand more information privately to the management of public sector agencies. While the production of the information may be useful to the ministers, the public disclosure of the information may not be useful to them. There may get sanctions from their peer (in the government and the Parliament) and the public for bad news in the performance information disclosure.

Further up in chain, the principals' interests are more diffuse and diverse. Pollit (2006) contends that the interests of the Parliamentary members is less likely to be intensive. The interest is much more episodic and strongly focused on bad news disclosed in the performance report. However, they are the decision makers for funding allocation in the public sector. Therefore, the performance information disclosure will be useful to help them in evaluating the public agencies' funding and spending proposals in the budget process.

At the top level of accountability chain, the interest is very diffuse and diverse. At the taxpayers' or citizens' level, the benefits of performance information disclosure are not easily identified. Taxpayers may use such information for their own local and particular purpose, such as getting health or education services (Pollitt, 2006). In this



small scope of purpose, the published performance information may benefit them. However, the benefit is very limited since the individual does not have an effective mechanism to provide immediate sanctions on publication of performance information. They either have to wait for the election time to vote out (discipline) their representatives, which is a highly imperfect control mechanism, or to move to other jurisdictions, which may be infeasible to many taxpayers since it involves very large personal costs (P. Smith, 1990).

## **B. Review of Prior Research**

This part of the thesis provides a review of the literature on the areas studied and the methodology used. Since the research focuses on performance reporting in the public sector, it is necessary to present first some published studies carried out by other researchers in this area. It is followed by an overview of the literature in the area of content analysis methodology and an outline of a disclosure index research instrument. This provides the basis for a discussion of the selected research instrument used later in this study.

### **1. Research in Public Sector Performance Reporting**

Studies focusing on performance reporting have tended to address local government rather than central government. The entities studied are varied in terms of their geographical location and the types of entities (see Appendix 1 for a list of some studies analysed). US-based studies have used cities as the object of the research (see for example, Ho & Ni, 2005; Poister & Streib, 1999; Robbins & Austin, 1986), while studies in other countries (such as UK, Australia and New Zealand) use local authorities (e.g., Boyne & Law, 1991, 2005; Ryan, Stanley & Nelson, 2002; Smith & Coy, 2000), schools (e.g., Tooley & Guthrie, 2007), police agencies (e.g., Collier, 2006) and even art galleries (e.g., Thompson, 1995). There have also been some studies published on local government reporting in Mediterranean and Scandinavian countries.

These studies, except for studies by Robbins and Austin (1986) and Ryan et al. (2002), which focus on the information quality of the annual report, provide some evidence on the status of performance reporting in the local government setting and the low quality of performance information. Performance information was viewed as generally poor and not improved over time (Boyne & Law, 1991) and to be the weakest section of the reports (Banks & Nelson, 1994).

On the other hand, the few studies focusing on performance reporting in a central government setting have mostly looked at performance reporting of UK executive agencies (Carter, 1991; Hyndman & Anderson, 1995, 1998; Hyndman & Eden, 2001, 2002; Pendlebury, Jones & Karbhari, 1994; Rutherford, 1996). The studies report on the type and the number of performance indicators reported by executive agencies. The limited findings show the increased provision of performance information (Carter, 1991; Rutherford, 1996) but mainly in terms of output information (Hyndman & Anderson, 1995, 1998; Pendlebury et al., 1994), and which is therefore inadequate for the discharge of accountability (Pendlebury et al., 1994; Rutherford, 1996).

Several studies use survey methods to find out whether there has been any improvement in performance reporting (e.g., Lawton, McKeivitt & Millar, 2000; Pollanen, 2005; Waweru, Porporato & Hoque, 2007). The findings support the notion that public officials prefer the use of effectiveness measures rather than efficiency measures, which is the opposite of the existing practice of disclosing more efficiency measures rather than effectiveness measures. The findings also show that they prefer measures that are internally developed (Lawton et al., 2000).

Other related studies investigate the degree of utilisation of performance information by the users as an indicator of information quality (e.g., Brusca Alijarde, 1997; Coy, Fisher & Gordon, 2001; Steccolini, 2004; Taylor & Rosair, 2000). Their findings are also disappointing. These studies have found that the number of users is limited (Coy et al., 2001), as the audience for the reports tends to be 'within-government' groups (Brusca Alijarde, 1997; Steccolini, 2004; Taylor & Rosair, 2000). These reports are intended more to satisfy accountability obligations (Taylor & Rosair, 2000) and less to communicate with external users (Steccolini, 2004). Performance information is not used by politicians and board members to control service delivery agents (Day & Klein, 1987). Even in Australia and New Zealand, where more extensive performance information is published, members of parliament make little use of it (Thomas, 1998). As Boyne and Law (1991) have noted, the findings reflect the poverty of performance information in annual reports.

Several other studies point out various important issues relating to the practice of performance information: obstacles in performance reporting, how the reporting dynamism develops over time, and the determinants of the quality of performance information. Studies that focus on the problems of developing and reporting performance information in the public sector have found several difficulties faced by agencies in the public sector. Stewart and Walsh (1994) found problems related to the

difficulty of determining a set of performance indicators suitable for services (which are the dominant outputs in the public sector) and the comparability of indicators across different agencies, activities and outputs. Rutherford (2000) found that public managers struggle to report the battery of performance indicators that provide a comprehensive picture of performance consistent with cost-benefit principles. Di Francesco (1999) found that public sector agencies were not doing well in the identification of direct links between outputs and outcomes (Di Francesco, 1999).

The presence of these problems means that development of performance information in a typical public sector institution over time has followed a distinct pattern (Mayston, 1985). Initially, information is produced in small batches of indicators from readily available data. This information has more curiosity value rather than functionality. Over time, the volume of indicator information increases, but there is no corresponding change in the organisation, design technology and functional intentions in the production of performance information. Thus, the volume increases but the 'quality' does not.

Two solutions are offered for these problems: regulation of information (Hyndman & Eden, 2002) and acquisition of skills and capability (Webber, 2004). Better systems can be developed more quickly if greater prescription and detailed guidance can be provided to regulate the performance information that needs to be reported. This situation calls for comprehensive and generally accepted standards for performance reporting. It also demands a well-specified performance measurement and performance reporting systems. The acquisition and development over time of the necessary skills and capabilities that enable organisations to effectively integrate the output and outcome concepts will lead to the eventual success of performance-based public sector reporting.

The research literature points to several factors that influence the type of performance indicators and the 'quality' of their disclosures by institutions. Carter (1991) offers seven dimensions of organisations that might be relevant to the shaping of performance indicators: ownership (private or public sector), trading status, level of competition, heterogeneity of outputs, political pressure, degree of complexity in output delivery, uncertainty in the objectives and between the means and the ends for achieving the objectives. Size of the organisation has also been found to have a significant effect on the quality of disclosure (Ryan et al., 2002; Taylor & Rosair, 2000).

Studies in the public sector reporting literature that focus on central government practice can be grouped under two perspectives. The first group of research, which is

the majority, is based loosely on rational choice theory. Within this group, most studies utilise some form of accountability concept to explain the role of reporting, such as studies by Hyndman & Anderson (1995, 1998), Hyndman & Eden (2001, 2002), Catusus & Grondlund (2005), Brun & Siegel (2006), and Herawaty & Hoque (2007). Some studies, for example research by Rutherford (2000), Carlin (2004), Brun & Siegel (2006), and Mack & Ryan (2006), have focused on the role and characteristics of information in the relationship between the agencies and their stakeholders. The second group of research uses an institutional and contextual perspective. Theories being used as the lens for analysis are stakeholder theory, such as research by Hoque et al. (2004) and Rantanen et al. (2007), and institutional theory, which has been used in research by Laegreid et al. (2006) and Hoque et al. (2004). The last group of studies combined these two perspectives as either a combination of theories, or a contrast of both theories, such as studies by Julnes & Holzer (2001), Laegreid et al. (2006), Sotirakou & Zeppou (2006), and Rantanen et al. (2007).

Although some studies based on some variation of rational choice theory exist in the public sector reporting literature, very few use agency theory or public choice theory as the lens for the analysis. In fact, even if the research is based on some idea from agency theory, all of them are just superficially based on agency theory, such as Mack and Ryan (2006). It is a disappointing situation, considering the public sector reform movement has been built on a foundation that incorporates both economic theories. A lack of research utilising these economic theories brings about a situation in which public sector practitioners are unable to assess whether the deficiencies of public sector reform under the umbrella of New Public Management are the result of inconsistent implementation of the economic theories or incompatibility of the theories with the public sector environment. Since agency theory and public choice theory are foundational for this reform, these theories should provide useful insights into the current situation, and therefore may be useful in revealing some of the limitations seen in reporting practice by comparing these with the theories. Agency theory may also be useful in identifying some of the limits in the current structure of contractual arrangements or the relationship between the principal and agent in the public sector environment. Public choice theory is useful for identifying whether the current structure is getting closer to implementing the “superior way” of producing public goods. These are the reasons why this thesis utilises agency theory and public choice theory as the framework for analysis.

## **2. Literature Review on Disclosure Measurement Methodology and Research**

To address one of the questions in this thesis, which is measuring the quality of performance information disclosed by New Zealand central government agencies, a measurement method/technique will be applied. There are a number of proxies used in the literature for measuring disclosure. Disclosure analysis in this thesis uses a variant of content analysis methodology, which is called the disclosure index. To provide a perspective on the methodology used, it is necessary to present a literature review of this methodology. The discussion will start with a summary of proxies used in disclosure research based on the work done by Hassan and Marston (2010). The summary will be followed by a description of the basic methodology in content analysis. The discussion will be concluded by a review of accounting disclosure research using a disclosure index.

### **Proxies for Disclosure in Accounting Research**

Hassan and Marston (2010) categorise proxies used in measuring the disclosure in accounting research into two groups. The first group consists of proxies that are not derived from a disclosure vehicle. The second group consists of proxies derived from a disclosure vehicle.

Proxies for disclosure in the first group are proxies that are generated externally from the disclosure vehicle. According to Hassan and Marston (2010), proxies in this group are either the indicator variable, which is a variable that indicates the level of disclosure, or the causal variable, which is a variable that causes the increase in the level of disclosure. Some examples of proxies in this group are as follows:

- The proxy as an indicator variable. The measure that indicates the level of disclosure is derived from the perception of analysts, investors or other user groups of the disclosure practice through the use of interviews or questionnaires. The advantages and the disadvantages of this approach depend on the interview or questionnaire instruments used. In addition, while the scores are not labour-intensive and can cover a sizable sample of organisations, there is a concern with the objectivity of the views of the user group investigated.
- The existence of the American Depository Receipt (ADR) is an indicator variable of disclosure level. Non-US firms that listed in the US market via the ADR mechanism are considered to have an increased level of disclosure. The

convergence with International Financial Reporting Standards has made this approach lose its usefulness as indicating increased disclosure.

- Attributes of analysts' forecast (AAF), and the number of analysts following the company are used as proxies for the indicator/causal variable for information disclosure. Firms with more informative disclosure are said to attract a larger analyst following and with more accurate forecasting. Therefore, the size of the analysts following and the accuracy of their forecasts are indicators of more informative disclosure. However, this notion has been challenged by the possibility of firms manipulating earnings toward financial analysts' forecasts.
- Other proxies such as period of listing and event (change of GAAP or reporting requirement) are used because of the limitations faced by researchers on time and access to relevant data. The use of the period of listing as a proxy is based on the logic that the longer the period a firm is listed in the market, the better is its disclosure practice. The use of an event of change in the reporting requirement is based on the assumption that the new requirement produces a better information environment.

Proxies for disclosure in the second group are proxies that are generated by analysing the disclosure vehicle directly. Some examples of proxies in this group are:

- Content analysis is a research technique that analyses the content of a text to make replicable and valid inferences from the text (Krippendorff, 1980). Hassan and Marston (2010) suggest that there are two types of content analysis: conceptual content analysis and relational content analysis. The conceptual type of analysis is used to determine the existence of concepts within a text by measuring the frequency of key words in the text. The relational type of analysis not only determines the existence of the concepts but also examines the relationships among concepts in a text.
- Disclosure index is a research technique that tries to measure the extent of information reported in a particular disclosure vehicle, according to a list of selected items of information that completely represent a certain theme. Hassan and Marston (2010) differentiate this technique from content analysis, while in this study, it is argued that this technique is a part of textual research techniques under the umbrella of content analysis. This technique will be discussed further below.

- Management forecasts are a specific type of information that may be disclosed in a disclosure vehicle. They are used as a proxy for disclosure quality due to the availability of verification of these forecasts from subsequent disclosure. However, similar to the problem of using analyst forecasts as proxies, these forecasts could also be subjected to earnings management, thus reducing their usefulness in research.
- The research may also classify firms' disclosure into good news or bad news and study the characteristics of good news in contrast to those of bad news.
- Disclosure frequency and changes in the frequency of disclosure are other proxies used to measure the level of disclosure. The level of disclosure is considered to increase if the disclosure frequency increases.

### **An Overview of Content Analysis**

The methodology used in earlier studies of accounting disclosure is based on the content analysis approach developed in communication studies. However, some aspects of content analysis have been modified specifically for accounting research and this results in the recognition of these approaches being dealt with separately from established content analysis methodology, by some researchers such as Hassan and Marston (2010).

Content analysis is a research methodology initially developed for research in the communication field of study to study the message of communication, not the communicator or the audience. As Kassirjian (1977) asserts, the unit of the analysis is the signs and symbols rather than the intent of the communicator or the actions of the audience. This analysis is an observational research method used to evaluate systematically the symbolic content of recorded information (Kolbe & Burnett, 1991). It is also a "scientific method that requires rigorous and systematic analysis" (Barcus 1959, as cited in (Holsti, 1968)).

The earliest definition of content analysis is offered by Berelson (1952), as quoted by Kassirjian (1977), stating that it is a research technique to describe the manifest content of communication in an objective, systematic, quantitative manner. Fearing (1954), as quoted by Kassirjian (1977), expands the definition by including latent content of communication as the object of the study and explaining that the process of analysis should be done by judges/scorers using objectively defined criteria. Thus, content analysis is a research method used to study the latent and manifest content

of recorded communication by employing judges to objectively and systematically apply categorisation rules based on selected criteria.

Accounting researchers have borrowed this method of analysis to use in their studies on narrative reporting. In the accounting literature, studies utilising content analysis are classified into two categories: studies investigating syntactic structure of a narrative and studies focusing on thematic structure of a narrative. The focus of syntactic analysis is the structural organisation of text, which is mainly about readability of narrative reporting in most accounting studies, while the focus of thematic analysis is the content of the information (Sydserff & Weetman, 2002).

Smith and Taffler (2000) further differentiate studies using thematic analysis into “form-oriented” analysis and “meaning-oriented” analysis. Form-oriented analysis is similar to the conceptual type of analysis according to Hassan and Marston (2010) that measures the frequency of key words or counting of words in a narrative reporting. In this analysis, researchers rely on some form of objective and computerised analysis of text based on a catalogue of keywords. While the counting itself is objective, there is a certain degree of subjectivity in constructing the catalogue of keywords. Meaning-oriented analysis focuses on identifying the underlying themes in the text under investigation. The themes can be constructed in the form of opposite themes, such as positive statement or negative statement, failure or success, internal or external attribution, or in the form of a list of themes that convey a complete (comprehensive) communication of a certain intention, such as reporting corporate social responsibility, reporting intangible assets, or reporting for accountability purposes. Since the themes identified can be constructed in several ways, researchers can then analyse whether there is a pattern of causal reasoning and attribution of a certain theme to another. For example, whether the theme that conveys success would relate to the theme that conveys internal attribution or whether the theme that conveys failure would relate to the theme that conveys external attribution. Alternatively, the themes identified can also be found in other organisational or external variables, such as the relationship between the comprehensiveness of disclosure in the area of corporate social responsibility and the degree of political pressure faced by firms. This second type of analysis covers the relational type of analysis of Hassan and Marston (2010) and disclosure index methodology.

Holsti (1968) claims that content analysis is useful for three general classes of research problems that may occur in most disciplines and areas of enquiry. First, content analysis will be important if researchers face data accessibility problems and the



available data is restricted to documentary evidence. Second, if the subject's own language and mode of expression is crucial in the investigation then content analysis is a necessary instrument for the investigation. Third, content analysis can be used in research in which materials should be evaluated objectively and systematically but the volume is too great to be handled by a single researcher. Content analysis allows for the use of trained assistants. This facility, in combination with systematic sampling of stimuli and interjudge reliability, makes content analysis an indispensable formal method of scientific analysis.

However, researchers need to be aware of the benefits and the limitations of this methodology. Kolbe and Burnett (1991) list several benefits and inherent weaknesses of content analysis. The benefits are:

- Content analysis permits an unobtrusive appraisal of communication, which can be valuable in situations in which biased responses are generated when researchers use direct methods of enquiry.
- Content analysis allows for relational analysis between environmental variables or source characteristics and message content.
- Content analysis can provide initial empirical evidence for further research about the nature and effect of specific communications.
- Content analysis is a useful companion research method in multi-method studies that can enhance the validity of the research results by mitigating method biases.

Kolbe and Burnett (1991) also list some inherent weaknesses of content analysis. The weaknesses are:

- Researcher biases can affect decisions made in the collection, analysis and interpretation of data.
- The potential of content analysis is limited to specific elements of communication. It is an exploratory research that is not anchored in specific theoretical perspectives. Therefore, it is not easy to identify the theoretical lenses used in the analysis.
- Content analysis in essence summarises complex communications as categorical data. Sensitivity to subtleties in communications is reduced in this method. Other methods with higher-order scales may provide better sensitivity to the nuances of communications.

There are three distinguishing characteristics of content analysis: objective, systematic and quantitative. The objective characteristic refers to the requirement that

the categories of analysis be defined so precisely that if different judges (analysts) apply them to the same text then the same results would be produced (Berelson, 1952). The subjective predispositions of the judges should be controlled at the minimum level possible, by requiring that all decisions made are guided by an explicit set of rules (Holsti, 1968). The main test for objectivity is that if other judges use identical procedures and the same set of data, they will arrive at similar conclusions (Kassarjian, 1977).

According to Kolbe and Burnett (1991), the objectivity of content analysis can be enhanced if research using this methodology includes two factors. First, there should be precise operational definitions and detailed rules and procedures for coding, to facilitate an accurate and reliable coding process. Second, multiple and independent judges are employed in the research and all of them are properly trained. Therefore, good research utilising content analysis should report five elements: rules and procedures, training of judges, pretesting of measures, independence of judges from the author and independence of judges from one another.

The systematisation characteristic refers to the requirement that analysis must be designed to obtain data relevant to the hypothesis or problem (Berelson, 1952). It means that criteria/rules used in the coding process should be based on the problem investigated or hypothesis stated in the research. The criteria should be consistently applied in the inclusion and exclusion of communication contents or analysis category (Holsti, 1968).

The last characteristic, quantification, refers to a measurement of the extent of emphasis or omission of any specified analytic category. The result of content analysis should be a quantitative score of qualitative data (text). Quantification can be in the form of a frequency count or a score to represent an emphasis or theme. With quantification, data resulting from content analysis can be amenable to statistical methods for precise and parsimonious findings and for interpretation and inference.

Kassarjian (1977) asserts that a typical procedure for content analysis should include four important steps. First, there should be a method of determining and selecting a reasonable-sized sample for study from the available population of documents to be studied. Second, there should be a determination of the units of measurement, such as specific words, an overall theme, or the existence or non-existence of an item in the criteria. Third, the judges should be trained for categorisation of the content according to predetermined rules. And the last is that the procedure should include treatment and analysis of data, including the use of statistical analysis.

There are several choices for measurement units in content analysis. The first one is words, which may include compound words. Readability studies commonly use the word unit in the analysis. The second choice is a theme, which is defined by Kassirjian (1977) as a single assertion about a subject. He further argues that this unit of measurement is the most useful unit of content analysis but it is also the most difficult one. The discussion of issues, values, beliefs and attitudes is mostly in this form, so it is very useful for meaning-oriented analysis. However, it is not easy to identify themes in a text. In some cases, more than one theme may be present in a sentence. The judge must be able to identify these component themes before putting them into the proper categories (Holsti, 1968). In other cases, a single theme may be present in several sentences or paragraphs or in a non-textual presentation. The third choice is characters. This unit is mostly used in studies of fiction, drama, movies, radio or any other form of entertainment material (Kassirjian, 1977). The fourth choice is item, which is the whole text or material of the analysis. In a typical piece of accounting research, this may include a financial report, a company release, a letter, or any single identifiable form of communication material. The last choice is the space and time unit. The measures for this unit can be column inch, the line, the paragraph, the minute, the page, or the foot (Kassirjian, 1977).

### **Reliability and Validity in Content Analysis**

Kassirjian (1977) argues that reliability, or reproducibility, is the strong point of content analysis compared with other techniques of analysing communication content. There are two types of reliability: category reliability and inter-judge or inter-coder reliability. The first type of reliability depends upon the ability of the researcher to formulate categories and to present definitions of categories so that competent judges will agree on identifying which unit of analysis belongs to a certain category and which does not. It is about whether the criteria can be used by the judges to identify easily which unit of analysis belongs to a certain category in the criteria and not to the other category. If the coding is done several times using the same material and the same criteria, there is a high degree of consistency in the results of these trials in criteria with strong category reliability. This type of reliability is also called stability, which is the weakest type of reliability (Krippendorff, 1980). A common measure for this reliability is the ratio of coding agreements from one trial to another to the total number of coding decisions. The test-retest is the procedure to assess this reliability. In this procedure, a judge will repeat the same analysis after an interval of time.

Inter-coder reliability is the measure of agreement among several judges coding the same set of research material using the same criteria. This reliability is the degree of consistency between judges. It is also called reproducibility (Krippendorff, 1980). This reliability is commonly measured by comparing the number of coding agreements with the total number of coding decisions. Since there is a possibility that some agreements may have occurred randomly, several different adjustments to this measure have been proposed, such as Krippendorff (1980)'s alpha and Cohen (1960)'s kappa. This reliability is influenced by the type of categories used and the decision made by the judges. Reliability can be increased if categories are narrowed and simplified and the decisions are made simple. However, the narrower the category, the less valuable is the validity of coding criteria (Holsti, 1968).

Milne and Adler (1999) argue that it is not an easy task to formally measure inter-coder reliability and to establish minimum standards to be achieved in disclosure studies. Each method developed using content analysis generates different theoretical and working limits. Therefore, there is no general rule that can be adopted in order to set an acceptable level of reliability in disclosure research using content analysis. They only advise on two aspects: researchers should be careful in interpreting the results of the analysis and researchers need to understand the tools, their limits and the research context.

Validity of a methodology is also an important issue. Validity refers to the extent to which an instrument measures what it claims to measure (Kassarjian, 1977). There are many types of validity, but the most important one for content analysis is face validity or content validity. Face validity is about the accuracy of correspondence between construct (such as theme, level of accountability, level of comprehensiveness) and its respective measurement (such as coding methods, disclosure index) (Jones & Shoemaker, 1994). Usually, this validity is assessed by seeking a subjective judgment or opinion from professionals or experts in the area that is measured by the instruments. For example, disclosure indices on intangible asset disclosure are usually discussed with experts in the reporting of intangible assets and some interest groups (such as managers, bankers and analysts).

### **Research on Disclosure using Disclosure Index Methodology**

The objective of thematic content analysis is to identify and analyse themes, such as specific trends, attitudes, or content categories, inherent in the narrative (Jones & Shoemaker, 1994). This type of content analysis has been widely used in the

accounting literature to study accounting narrative, especially in studies of the extent of disclosure in annual reports.

Guthrie & Abeysekera (2006) categorise content analysis used in measuring disclosure into two approaches. The first approach is to identify the overriding themes in a text by way of measuring the frequency of words or expressions that resemble the themes. This approach is usually used in research that investigates the presence of opposite themes, as previously explained. The second approach uses disclosure index as the research instrument. The focus of the first approach is on the quantity of disclosure. The disclosed information is measured by counting all the data items, such as number of words and numbers. The problem with this approach is there is repetition of certain numbers and words in the accounting narrative. In addition, numbers on their own do not have informational content or represent a single category unless they are accompanied by explanatory words (Marston & Shrives, 1991).

Studies focusing on quality of disclosure assess what is actually being disclosed. As Guthrie and Parker (1990) assert, these studies should focus on what was said and how it was said: theme, evidence (monetary, non-monetary, declarative, none), amount and location of a disclosure. Similarly, in their study on the quality of environmental disclosure, Gray et al. (1995) and Hackston and Milne (1996) focus on examining themes, evidences, amount, auditability and news to infer the quality of disclosure.

Disclosure index is the selected instrument used in most studies focusing on the quality of disclosure. This instrument, as stated by Marston and Shrives (1991), is a list of themes or disclosure categories to assess, compare and explain differences in the extent or comprehensiveness of disclosure. There is a well-established area of disclosure study utilising this instrument in the accounting literature. A majority of research focuses on whether disclosure of specific items is considered important for accountability. An accountability index or modified accountability index is quite popular in this area of the accounting literature.

A disclosure index is a research instrument used to measure the comprehensiveness or the level of disclosure in the specific context for which the index is devised (Guthrie & Abeysekera, 2006). It contains a series of preselected items, in which each will be scored based on the existence of the item in the targeted documents or the extent to which the information on that item is disclosed. A comprehensive disclosure for a certain context means that all relevant information for that context is reported. For example, an annual report of a public institution is considered comprehensive for accountability purposes if all relevant financial and non-financial

information has been reliably reported. The researcher then develops a public accountability index containing a list of items that should be reported in the annual report that is considered to fulfil the accountability purpose. When the index is used to assess annual reports of public institutions, the score result is viewed as the level of comprehensiveness of the reports to fulfil their purpose of delivering accountability.

The selection of the form of disclosure index involves a trade-off between reliability and validity. The simplest form of disclosure index uses a binary coding system. Each item measures the presence or omission of information in the assessed documents. The total scores represent an aggregate measure of the quantity of disclosure. This form of index usually produces the highest level of reliability if it is designed properly by the researcher. However, unless the level of quality can be represented by some form of quantity or comprehensiveness measure, this index has lower validity than more complex indices. Complex indices allow for variation in the quality of individual items. The quality of disclosure for each item can be assessed using ordinal scales. Thus, validity increases if the quality of reported information is measured in a certain context. In spite of this, it is not easy to design a non-binary scoring system that has a high degree of reliability. More judgment is needed in scoring with ordinal scales and this reduces the reliability of the index.

There have been numerous studies on annual report disclosure. Some of the research has concentrated on overall annual report disclosures, such as research by Adams & Hossain (1998), Baker & Haslam (1973), Buzby (1974), Choi (1973), Chow & Wong-Boren (1987), Leventis & Weetman (2004), Singhvi & Desai (1971) and Zarzeski (1996). Other studies focused on specific information in the area of social corporate responsibility, environmental reporting and intellectual capital, such as studies by Adams et al. (1998), Bozollan et al. (2003), Brennan (2001), Bukh et al. (2005), Cowan & Gadenne (2005), Guthrie & Abeysekera (2006), Ho & Taylor (2007), Ingram & Frazier (1980), Milne & Adler (1999), Ratanajongkol et al. (2006), Roberts (1992) and Vandemaele et al. (2005). Some other studies focus on the disclosure of specific items under the notion of public accountability, for example studies by Bank et al. (1997), Gray & Haslam (1990), Hooks et al. (2001) and Ryan et al. (2002).

A disclosure index is a popular approach used in accounting research that investigates the quality of information issued to the public. From the sample of studies identified that use a disclosure index (see Appendix 2 for the list), most of the studies take all the information in the annual report as the object of analysis (Belkaoui & Karpik, 1989; Brennan, 2001; Cowen, Ferreri & Parker, 1987; Leventis & Weetman,

2004; Robbins & Austin, 1986). These studies investigate the relationship between external/internal variables of an organisation and the quality or extent of information disclosure. One study examined the quality of information in a prospectus (Bukh et al., 2005), and its relationship with organisational variables. Another study (Bryan, 1997) investigated a section in the annual report (management discussion and analysis) while another study (Ho & Taylor, 2007) investigated the annual reports, stand-alone reports and special website reports for similar associations.

The indices used in these studies are based on a variety of sources. The bases cited are literature review (Bukh et al., 2005; Ho & Taylor, 2007; Leventis & Weetman, 2004), surveys carried out by others (Belkaoui & Karpik, 1989; Cowen et al., 1987), surveys carried out by the researchers (Robbins & Austin, 1986), frameworks developed by other researchers (Brennan, 2001; Leventis & Weetman, 2004), and disclosure regulation (Bryan, 1997). Most of the studies, other than Robbins and Austin (1986), use unweighted scores. One study (Bryan, 1997) looked at the direction of disclosure (unfavourable, neutral, favourable or missing) instead of just the presence or omission of the information.

## **IV. METHODOLOGY**

### **A. Overview**

This study utilised library research, archival data analysis and semi-structured interviews. The library research was conducted to develop the possible and relevant theoretical arguments in the practice and quality of outcome reporting in the public sector, as well as the possible determinants of the quality of the reporting. Disclosure analysis using an index was used to assess the current practice of output and outcome disclosure over the period of analysis, in the form of a score. To determine the internal and external factors that influenced the increase in the quality of reporting, a semi-structured interview, combined with archival data analysis as far as possible to validate the interview results, was the main method used. The results were then compared and contrasted with the theory selected.

### **B. Research Method**

Three critical steps were carried out in the research for this study. First was the development of a disclosure index as the measurement tool. Second was the application of the index to score the comprehensiveness of performance information in the SOIs and annual reports. Third was conducting interviews and analysing the results in conjunction with the results of the disclosure analysis, to explain the factors that influenced the observed performance reporting practice.

#### **1. Disclosure Analysis**

The methodology used in this research can loosely be described as the meaning-orientated (subjective) approach of thematic content analysis. This approach focuses on the quality of the content to represent certain themes. What is being disclosed is not indicated solely by the quantity of the content (Frost & Wilmshurst, 2000). Some studies in this approach focus on the investigation of the presence or omission of certain themes (Sydserff & Weetman, 2002). Other studies in this approach require researchers to select (develop) narrative (disclosure) indices to assess, compare and explain differences in the extent and comprehensiveness of the content in accounting narratives (Marston & Shriver, 1991). For example, a researcher could study whether the disclosure practices comply with regulation and/or the level and extent of voluntary



disclosure. An index could be developed to include a mixture of required and voluntary items to suit the purpose of the research. This disclosure index method is the one which was used in this research.

The decision to use disclosure index was based on the fact that there is no other proxy available that represents the quality or level of comprehensiveness of performance information. While the overall assessment of information quality in the New Zealand public sector, as published by the Office of Auditor General (2008), is available, there is no publicly available assessment result for each individual public agency in New Zealand. Therefore, there is a need to develop and apply a proxy to represent the quality of information to permit this individual assessment. In addition, the study also offered an assessment method of information that had some degree of objectivity and which could provide quantitative assessment for qualitative objects.

The disclosure index technique was the method that was deemed appropriate since it offered a way to answer some of the problems addressed in this study. First, the quality of performance information could be measured with this technique. Second, the scores produced by applying this technique could be used to find out whether there had been progress in the reporting of outcome information by NZ central government agencies over time. Third, the scores could also indicate the extent to which the disclosure of performance information by NZ central government agencies had followed the available guidance or benchmark over time.

Methodological contribution of this study is the application of the disclosure index technique to disclosure research on non-financial performance reporting in the public sector. The disclosure index technique has not previously been used in studies that focus on non-financial performance reporting. The common technique to assess the quality of performance information is by indicator counting and categorisation. Disclosure index is usually used in public sector studies that assess the quality of overall disclosure in relation to the fulfilment of accountability objectives. These studies use some variants of disclosure index called a public accountability index. Whilst studies on specific reporting issues in public sector, such as those reporting on intangible assets or environmental aspects of operation, have utilised some versions of disclosure index techniques, currently there is no single research study in the public sector that applies this technique to investigate the quality of specific parts of the information disclosed. Thus, a key aim of this study was to fill this gap by demonstrating that disclosure index is a powerful yet flexible measurement technique that can be used in any research that attempts to assess the quality of a narrative.

## **Research Index Design**

Before designing a disclosure index, researchers need to specify the purpose of this instrument in their research. Once the purpose has been decided, items to be included in the index are selected. The index will then consist of a series of pre-selected items with a scoring system that can show the level of disclosure in the context of the purpose of the study. Therefore, the usefulness of the index as a measurement tool critically depends on the selection of items to be included in the index and the related scoring system.

The purpose of this research was to assess the changes in comprehensiveness of output and outcome information over a period of time; therefore the items included had to consist of an ideal or a complete set of output and outcome information. This set was approached by investigating the current literature and guidance related to the targeted information. Since currently there is no accepted framework for output and outcome information, documents that are available as guidance and reference for reporting outputs and outcomes, as previously outlined at the end of Chapter II of this thesis (see page 46-48 above), were used to identify items forming an ideal set of performance information.

When the item was identified, the possible level of disclosure of each item for scoring development was determined. There are four distinct levels of measurement: the binary scale, the ordinal or ranking scale, the interval scale and the ratio scale. Since the last two scales are not applicable in disclosure index methodology, the ordinal scale as a higher order scale was considered first. However, after the first iteration of the scoring process using this scale, the reliability of this scoring was not at an acceptable level. When the scoring system was trialled by different scorers (judges), they came up with quite different score results and disagreements between scorers were more the rule than the exception. The scorers had to employ a high degree of judgment in considering a score from more than one scale. While it was easy to spot the zero level of disclosure on the items, it was quite difficult to agree whether a certain disclosure should be assessed at the low level (score 1), medium level (score 2) or high level (score 3). The level of scoring consistency was not at an acceptable level, so it was decided that a binary scale instead would be used in this study. Even though binary coding is the least preferred method of capturing an increase in the comprehensiveness of information, it was used to get more objective and consistent scoring. The comprehensiveness was captured by the number of items reported, or in this case the number of total scores, and it was considered as being the proxy for information comprehensiveness.

The selection of binary coding made it easier to score the items; however it required an expanded set of items of disclosure. The previous list of items was expanded to capture more detailed components of disclosure than was previously a part of an item. For example, an item that previously was scored on a three-rank scale (0, 1, 2, and 3, with a 0 score for no information disclosed) was then expanded into three items, each with binary scaling (0 and 1). It was easier for the scorers, since they simply needed to identify whether the item was disclosed in the documents. If the information was presented, the scorers assigned a score of 1 for that item category in the index.

The development of disclosure indices provided some important lessons. The index should be neutral to the number of outputs reported. Some agencies may report many outputs and some may disclose a few outputs. The score should not be affected by this difference. The scoring system should be able to make the same differentiation for binary-type outputs (especially for policy advice) and for high-volume outputs. In the case of binary-type outputs, the development of items to represent a more comprehensive disclosure was implemented after scanning a sample of documents to identify the variation of how the information on these outputs was presented. This type of output created something of a challenge because when it was reported, the information about planned and actual quantity was automatically presented, since its quantity was one. In high-volume outputs, this information needs to be presented separately. The results of the document scan showed that there were two variations on binary-type output reporting: a simple presentation only disclosing the name of the output and a detailed presentation disclosing not only the name of the output but also the milestones or deliverables of the output. It was decided that for this type of output, the information of planned and actual quantity was presented if the document contained planned and actual milestones or deliverables for the outputs.

Clear instructions were developed for scoring the comprehensiveness of performance information of binary-type outputs. Decisions also needed to be made in cases where the targeted indicators were not explicitly noted. For example, some indicators of actual results were presented in percentage form. A consistent rule was needed to decide whether the targeted indicator was simply not available or the target was implicitly stated at 100%. The most important lesson learned was that the scoring instructions should be as clear as possible, to enable consistent treatment across the sample.

The scoring system also had to take into account the fact that different documents assessed have different applicable sets of disclosure items. For example,

some documents may contain variation in the measures reported, and therefore a disclosure item that disclosed an explanation for this variation was expected. Consequently, for these documents the number of applicable disclosure items in their complete set of information was more than the other documents that did not report variation. In this case, index scores were later converted by dividing the actual score by the maximum score possible for every entity, to a relative score, as suggested by Marston and Shriver (1991).

### **Testing the Index**

Once the items were identified and the scoring method was designed, the draft of the index was developed, tested, validated and retested again. It was during this first test of the index draft that the decision was made to change the ranking scale into the binary scale, after considering the unacceptable level of consistency when using the ranking scale. The index based on the binary scale was then tested again. After an acceptable level of reliability was achieved, the index, which was called the alpha index, was ready to be tested using a bigger and more varied sample. The index was then tested to score a sample of reports. The reports were randomly selected from the population of documents to be scored. The purpose of this test was to find the degree of difficulty and the weaknesses of the scoring system that could reduce the reliability of the measurement. The researcher and an assistant familiar with government reporting in New Zealand performed the tests.

An auditor working at the OAG was an ideal candidate as a scorer/judge because of his extensive knowledge of public sector reporting in New Zealand through his experience as an auditor. Before he conducted the scoring of all documents assigned to him, he was trained in the disclosure index methodology. Then, the researcher and the assistant did a run-down of scoring instructions using a set of documents to be assessed (a pair, the Statement of Intent and the annual report). Any confusing instructions were clarified during the process. The assistant was then deemed to be ready to perform the scoring test for the alpha index.

This test revealed another difficulty in scoring the information, as most agencies only disclosed clear identification of performance indicators at output class level, instead of at a more detailed output level. The scoring at output level produced a higher degree of differentiation than the scoring at output class level. Therefore, it was later decided that the scoring would be applied at output level for some disclosure items. However, it was noted that there was a substantial degree of inconsistency in identifying

the outputs in each output class for some agencies, especially those that provided no clear identification of outputs within their output classes. To overcome the difficulty, it was decided that the output identification would be done separately. Before conducting an assessment of a set of documents, each scorer identified the outputs individually and met later to resolve the difference.

Once the scores agreed on the output identification, they proceeded with the full assessment of the documents. The scorers conducted the scoring separately for the sample reports. The test results between these two for the same sets of documents were then compared so that the reliability of the scoring system could be assessed. The scoring system proved to have a high degree of consistency, with less than a 1% scoring difference between the two scorers. The difference was discussed and the scoring system was refined further.

### **Validating the Index**

The next step was presenting the index in two rounds to a group of selected experts in the New Zealand public sector reporting, for four purposes: (1) validating the items proposed in the index, (2) capturing any missing items suggested by the experts, (3) assigning weights to the items in the index, (4) assessing the validity as well as reliability of the scoring system. The experts were academicians in public sector accounting, officials from the OAG and a former official of the Treasury who was previously involved with the MfO initiative. The validated index used in the final scoring, as shown in Appendix A, contains the expert corrections and recommendations. The index weighting was abandoned, as some of the experts were concerned that applying the weighting would distort the index from its intention of capturing the comprehensiveness of the information. Once the index was validated, it was retested again against a very small sample of reports for the same purpose as the test conducted in the alpha index, but directed mainly to new items and/or a new scoring system resulting from the consultations. At this stage, the scoring was also done by two people for the purpose of calculating the reliability measure.

### **Reliability Testing**

Reliability of the methodology was the main concern in performing the disclosure analysis. The index score awarded to agencies was considered reliable if the results were consistent from one assessment to another and, when other researchers repeated the process, the results were similar. Since documents in which the scores were

extracted from remain constant over time, there will be no threat to reliability from the data source. However, the researcher applying the index and the scoring system to award the score had some potential problems with the reliability of the research. One attempt to increase the reliability was to have the level of subjectivity in scoring kept at a minimum level by providing clearer instructions for some problematic areas identified during the testing stage. Later, two reliability tests were conducted. The first was a test-retest procedure. The researcher scored all the reports and redid the scoring for a randomly selected sample of the reports one month later. The scoring results from the same sample of the two tests were compared and the level of disagreement was less than 1%. The research assistant, acting as an 'expert' judge, also did the scoring for a randomly selected sample of the reports. For reasons of resource limitation, the second scorer (the assistant) only scored 25 pairs of documents (50 documents) that were randomly selected. The scoring results from the same sample of the two judges were compared and the index was deemed to be reliable, as the level of disagreement was less than 1%.

### **Data Set and Disclosure Analysis**

The periods covered in this research were the first five of the fiscal years affected by the MfO initiative. The ex-ante reports (SOI) and the related annual reports for the period of 2003/2004 to 2007/2008 were used, and therefore for each executive agency 10 documents were analysed. Most of the reports were downloaded from the agencies' websites. Since some agencies could not provide all the reports, the reports were requested by applying directly to the agencies. Some agencies were new, or as a result of reorganisation they did not have the complete set of documents for the period covered in this analysis. Therefore, they were not selected for this analysis. Some other agencies could not provide all the required reports for the period covered, so they were also excluded. Out of 41 agencies, 27 agencies were selected as a sample for this research, because all documents for the analysis were available for these entities. In total, 270 documents were analysed in this study. The list of the agencies selected is available in Appendix D.

To assess the representativeness of the sample, the list of departments/ministers was referred to three academicians with extensive experience of the New Zealand public sector. None of them raised any concern about the selection being biased. They concluded that the sample was representative of the population of New Zealand central government agencies.

After all documents were scored, the overall results were analysed to find out whether some components of the index had such a low occurrence level that the inclusion of them would skew the overall results. In this analysis, it was decided that one disclosure item group, which is location performance information, should be dropped from the list and the scoring results, since only three agencies disclosed this information, and two of them only disclosed it for one set of accountability documents.

The total raw scores for each department were then transformed into relative scores, to take into account the fact that some agencies may have more disclosure items due to the existence of variances in their reported measures. This relative score of an agency was calculated by dividing the raw score and the number of total disclosure items applicable to that agency. In essence, this relative score represents the percentage of performance information disclosed in relation to the complete set of that performance information.

### **Limitations**

The validity of this study's finding may be subject to the limitations of the disclosure index technique. First, there was an assumption that the scope (or coverage) of a disclosure was a proxy for disclosure quality. The validity of the scoring results to represent the quality of the disclosure depended on the acceptance of this assumption. Second, the method involved scoring and quantification to come up with a single figure representing the overall quality of a narrative. It may take into account the relative importance of each item in the index by way of using weighted scores. The study did not take this into account due to the consideration that the relative importance of each item may be cancelled out if the narrative is intended to be a means of communication to the general public. One factor that is deemed important by some members of the public may be considered as not important by other members. Unless a study is intended to assess the quality of information from the perspective of a specific group, un-weighted scores are more appropriate. However, this consideration has not been subject to a rigorous analysis. Third, construct validity cannot be assessed since this was the first study that assessed the quality of non-financial performance reporting using a disclosure index technique. Fourth, there was a level of subjectivity in the selection of the items and the coding process. Therefore, it was only valid to the extent that the index used was appropriate (Hassan and Marston, 2010).

## **2. Case Study for Analysis of Relevant Factors**

To identify the external and internal factors influencing reporting practice, a case study was utilised. This methodology was selected because it facilitated the understanding of complex inter-relationships in the public sector relating to performance reporting practice. It could also facilitate the exploration of the process involved in the causal relationships. This methodology fitted with the purpose of the research in describing complex factors that were likely to influence reporting practice by investigating the complex relationships of actors in the public sector.

There are five steps to take in a typical case study. The first step is to determine the theoretical framework used to develop the pattern of concepts to be identified in the case study. This framework provides some notions of where to begin the inquiry. It also serves as a sort of checklist of points to investigate during the course of the study. The next step is to determine and apply appropriate data collection methods. The objective of this step is to obtain rich data that can be used in the analysis. Case studies usually use multiple data collection techniques, such as observation, document study, interview, documentation etc. This study used only two methods of data collection, which were interview and document study. Once data collection was completed, the analysis was carried out to identify and tabulate meaningful events as directed by the research framework. These events serve as findings in the research to support the next step of building explanation. In this step, explanations are developed by comparing and contrasting the findings with the concept pattern in the theoretical framework in order to confirm or reject all or parts of the concepts. The last step is a consideration for conducting multiple case studies. In this study, this consideration was important because the case studies undertaken were not intended to be used to gain an understanding of a particular issue pertaining to any particular entity. The aim was to investigate the factors that even common to a certain group of entities. Thus, case comparison was important to ensure that the findings could be generalised to a certain extent.

This study used multiple cases to generate findings. The same case study methodology was repeated with a sample of entities so that similar results could be identified and contrasting results could be explained. The strength of this multiple case study was that it permitted some level of generalisation. However, in order to cover more cases, the depth of the study was reduced so that the investigation of multiple cases was still manageable. Consequently, the depth of analysis in this study may not be as deep as the analysis in a typical case study.



Data was collected in this study using interview and document study. Observation technique was not used because the events investigated occurred in the past. Out of these two techniques, interview was the main means for data collection since most of the concepts or events being investigated were mainly in the form of past experience. Interviews are particularly useful in obtaining the story behind a participant's experiences. Interview can also be used to pursue in-depth information on a topic. When relevant documents are available, they are analysed to gather findings that can be compared and contrasted with the results of the interviews. Interviews may provide an insight into the mechanisms of implementation and the causal links to a particular programme or initiative. It helps to identify success stories and obvious shortcomings. It may also help in proposing solutions and recommendations for initiatives in the future.

### **Identification of Interview Targets and Initial Contact**

The agencies were grouped into three categories based on their disclosure scores. The first group consisted of nine agencies with the highest scores, the second group consisted of those with the next nine highest scores, and the last group consisted of the nine agencies with the lowest scores. The rank was determined using the average of total scores for all periods covered. Three agencies were selected randomly from each group. They were subject to two requirements: first, the agencies were not central agencies, and second, the official(s) that prepared the accountability reports at least worked with the agencies during the last three years of the period covered in this study. The focus of the case studies was those agencies that had a pure role as the preparers of information. The central agencies were not selected, because they have an additional role in the MfO initiative as the regulator. Since interviews were targeted at the officials who understood the process and the development of accountability documents during the period covered in this study, it was necessary to target the agencies that still employed those officials. Five agencies that were selected said in their response to the interview request that none of their officials met the second requirement mentioned above. Therefore, other agencies were selected to replace these agencies.

The central agencies (in this case the Treasury and the SSC) and the OAG were targeted for a different role, not as preparers of the information but in their role as the regulators and as supporters of the MfO initiative. These agencies were also consulted to corroborate some of the issues raised by the preparers, in relation to the level of support in the MfO initiative and the regulation of performance reporting.

## **Interview Questionnaire Development**

To identify factors that promote or hinder improvements in the performance reporting practice of New Zealand central agencies, a framework of analysis was developed to guide the process of data collection and data analysis. Three aspects in the agency theory framework were investigated in this analysis. First, to investigate the existence of information asymmetry, the roles and actions of ministers (as principals) and chief executives (as agents) in the development of performance measures and targets in the planning process, especially those that were related to the provision of information, were explored. Document studies and interviews with officials from selected departments and the Treasury were the main methods of data collection. Second, to identify problems, constraints and difficulties, data was gathered from interviews with officials of the selected departments, the Treasury and the OAG. Third, these interviews were also a means for collecting data used to analyse various mechanisms used or operated by the principals, via the Treasury, the SSC, and the OAG, and to assess the effectiveness of these mechanisms in improving the quality of performance information over time. These interviews, to the extent that it was possible, were corroborated with the data by studying publicly available documents.

The list of questions asked in the interviews was developed by referencing the theoretical framework used, which in this case was a combination of public choice theory and agency theory. The draft of the questions was tested three times with three former public officials in New Zealand that were involved or knowledgeable in the area of output and outcome reporting. The tests were conducted to ensure that the questions were appropriate for the New Zealand public sector context and the concepts used were easily understood by targeted respondents. The tests were intended to identify any misunderstandings that might occur on the part of the interviewees. The list of the questions was later revised to reflect recommendations from the tests. This list is presented in Appendix I.

## **Interview Process and Data Analysis**

In the end, while 15 interviews were planned only 14 interviews were able to be conducted for this study. One was cancelled due to a scheduling issue, but the interviewees provided a joint written reply on the list of questions sent to them. Interviewees were from two central agencies (SSC and the Treasury), the Office of the Auditor General, and nine other central government agencies. For interviews targeted to

agencies as preparers, two persons were interviewed from each agency: one was the top-level official responsible for the coordination of publication of accountability documents and the other was the lower-level official who was responsible for compiling information for accountability documents. In three cases, only one person was interviewed. In two out of these three cases, a scheduling problem was a factor, but the persons interviewed, who were lower-level staff, had informed the interviewer that they had discussed the question with their superiors and were able to speak on behalf of them on some of the questions asked. In another case, the person interviewed was from one small department where he was the only person responsible for the whole process of coordination and development of accountability documents. In the other six agencies, two persons from each agency were interviewed. In four cases, the two persons were interviewed together. In two cases, because of a scheduling issue, the two persons were interviewed individually. From central agencies, three persons interviewed were high-level officials who were involved either in the MfO initiative or in the review of accountability documents. Two high-level officials from the Office of the Auditor General were interviewed together.

The main objective of this process was to get information that was not available from archival data regarding the agencies' experience in the production of accountability documents and the challenges that they faced during the process. The other objective was to reveal the dynamic interactions between the agencies and their political principals on one side and between the agencies, central agencies and the auditor on the other. The questions asked were about the process of preparing output and outcome information, the interactions of agencies with various parties in relation to output and outcome information, the incentives for reporting performance information and the challenges experienced in the process.

The interviews were flexible and were done interactively. The order of the questions asked was different from one interview to another, depending on the issue that emerged first during the opening discussion. Some interviewees discussed some of the questions at length, while there was not much discussion of the same questions in other interviews. Therefore, while the set of questions being asked was the same for the nine agencies that were targeted as the preparers, the responses varied depending on the issues being stressed by each interviewee. Each interview lasted between three quarters of an hour and an hour and a quarter, except for one interview that lasted for two hours and a quarter due to the interviewee's desire to provide more elaborate responses for each question asked. The interviews were recorded and transcribed into Microsoft

Word. All interviewees declined when they were asked whether they would like to check the interview transcript, but some of them said they would like to receive the summary of findings from the interviews. This summary was sent to them and no objections were received in their responses.

The interviews were analysed according to the themes identified in the theoretical framework. For each theme, a code was created. The codes were assigned to relevant passages in the transcription text. The passages were then categorised in a tabular format based on the theme, the agency interviewed and the grouping of the agency. Four Word documents were created for this categorisation, three documents for these nine agencies and one document to categorise the relevant passages from central agencies' and the OAG's interviews. Each theme was then analysed to find similarities or dominant patterns to be identified as findings. Variations were noted and presented as a part of the findings.

### **Limitations**

The validity of the findings from case studies relates to the limitations of the methodology. The first limitation is that the findings cannot be used to make scientific generalisations. The findings can only be generalised in relation to the theoretical propositions used and to the selected sample investigated. The second limitation is that sample selection, data handling and analysis are prone to influence from a researcher bias. The cases (entities) selected are not representative of diverse populations, therefore the findings are not applicable to such populations. The identification of events or concepts from the document study and interview involves some degree of researcher's subjectivity. Therefore, the study cannot be replicated to come up with similar results. These case studies can only make tentative conclusions on how much a particular variable affects the outcome in a particular case or type of cases. The case studies are meant to be used to identify the scope of the influence of particular theories and to assess arguments about causal necessity or sufficiency in particular cases. Lastly, the limitation of this study is also related to the use of interviews as the data collection method. Interviewees may be reluctant to give truthful answers and instead provide rather more socially acceptable ones.

## **V. FINDINGS AND ANALYSIS**

### **A. Findings and Analysis from Disclosure analysis**

The findings support the Auditor General (2008)'s assertions about the disappointing level of the quality of performance information in accountability documents. In addition, the study documented the pattern of increase in information comprehensiveness over time. In terms of output information, the level of information provided was predicted to be stable over the period of analysis, as the practice of disclosing this information had been established for more than a decade before the period of analysis. However, the only measure in the output information that was disclosed comprehensively was the quantity measure. The quality, timeliness, location and cost measures of outputs were only partially disclosed. In terms of outcome information, there was a noticeable jump in the amount of outcome information provided in the first period, but a very small increase in information over time. On average, the information disclosed was about 24% of the complete set of information in the first year of the MfO initiative, increasing to only 34% at the end of the period under analysis. The quality of disclosure was good in only two aspects: the outcome definition and the outcome framework. Other aspects, such as outcome measures, intervention logic (attribution from output to outcome) and risk identification were partially or poorly disclosed.

#### **1. Output Information**

The quality of output information in the accountability document tended to be stable over time, as presented in Table 1 of Appendix E. The average relative score remained stable over time, around 0.68 – 0.69. This meant that the comprehensiveness of information reported was on average 68% of the complete set of information suggested by reporting guidance. The yearly scores for each agency also did not vary much. This reflected a high level of reporting consistency over time. This result was expected, since the reported information was a product of the first established regime of New Zealand public sector reforms during the late 1980s. Since the reporting practice was already well established, there was a definite structure found in the reporting environment that meant the reported output information was consistently presented over time. In addition, the output information was reported in a specific format in the

Statement of Service Performance and was audited, so there were strong incentives for consistency in reporting.

The results presented in Table 1 of Appendix E also show that there was considerable variation in quality among the agencies. The best agency provided 86% of the complete set of information, while the worst agency provided only 27%. The New Zealand Defence Force, an agency with a service delivery focus, set a good example with its reporting practice, while the worst agency was the Ministry for the Environment, which focused on policy advice delivery. However, there was no indication that policy advice agencies produced lower quality performance information than service delivery agencies. Three of the top five agencies in terms of output information quality were policy advice agencies and three of the bottom five agencies were service delivery agencies. The high degree of cross-sectional variation in quality reflected the fact that there was still a high degree of discretion in reporting output information. It seems that there was no mechanism to ensure that cross-sectional consistency was promoted. It is also important to note that when interviewing the respondents from some of these bottom five agencies, they acknowledged that they had been given an unfavourable assessment in a management letter from the auditor regarding the quality of output information. However, they did receive a clean audit report for the information.

The agencies presented output information at a high level in the areas of information that included output description, quantity information, quality information and timeliness information, as can be seen in Table 2 of Appendix E. All documents analysed disclosed all output descriptions at output class level, which was the level of detail of output information included in the budget process, but around 7% did not disclose this information at a more detailed level for each individual output. One disclosure item that was not reported at all well in the output description item group was the explanation of output class changes. TPA-9 requires the preparer to provide a description of output class changes, but for the majority of output class changes, no explanation was disclosed. Agencies also provided almost all cost information at the output class level. It seems that if the information was available via a budget process, this information was reported completely in the accountability documents, resulting in almost perfect scores.

The scoring for information on quantity, quality and timeliness was assessed at a detailed level for each individual output, so it was more refined and showed more differentiation when the quality of information was assessed. In these groups of

disclosure items, there are a few things to be noted. First, agencies did not disclose performance measures for all outputs. The most comprehensive one was quantity information, followed by quality and timeliness information. Second, some planned or targeted measures were not reported for these three types of performance measures. Third, there was no consistency rule with reporting variances. In some cases small variances were reported, in other cases the larger ones were not reported. To ensure consistent scoring, it was determined that if the actual result differed by more than 10% from the planned results, it was considered a variance for which an explanation should be presented. Therefore, the scoring results on these variance disclosure items should be interpreted as the percentage of variance information reported consistently under this rule. The 10% threshold was selected, since it was the figure used the most as a threshold in the accountability documents analysed that did report variance information.

The most disappointing result was seen in the information quality of cost measures at output level and also disclosure on the cost allocation method. The cost allocation method was disclosed by only one agency in one reporting period (the New Zealand Police for the accountability documents for 2004/2005). Less than 20% of accountability documents analysed disclosed cost information at the output level. One possible explanation for the low quality found in this area of information was because the TPA-9 only recommends this information to be disclosed for encouraging transparency in government. Since it was voluntary information, it seems that agencies would report it if it was already available within their systems. It might also indicate that the majority of agencies were not managed efficiently at output level.

One important thing that should be noted here is that it seems that the level of comprehensiveness of output information has not increased significantly, even during the last 15 years. A carried out done by Coopers and Lybrand in 1995 on departmental and Crown financial reporting, as quoted by Petrie and Webber (2006), reported the same result. The reviewers concluded that the output information was quite accurate at output class level. However, there were strong concerns about the system and the accuracy of performance information at output level.

To provide a more comprehensive picture, a detailed analysis of the score results is presented below. The discussion is divided into five parts: output description, output costs, quantity dimension, quality dimension, and timeliness dimension.

## **Output Description**

There are four items in this group. These items are information at the output class level: description of output classes, a list of detailed output, a description of detailed output, and explanation of output class changes.

As previously explained, performance information at output classes was more comprehensively reported by agencies. All agencies described their output classes in the accountability documents. However, there was one item of information at this level in which the scores were disappointing. Agencies are required by the TPA-9 to explain if there is an output class change. The average score for this item for the five-year period was 33%. Thus, the agencies properly explained it only once in every three cases of output class changes. There is a possibility that the disclosure of this information was ignored because output class changes were not very frequent. Nevertheless, this information should be reported, considering that it is information required to be disclosed if TPA-9 is followed properly. The omission of this information may be due to the fact that TPA-9 was not referred to in the reporting of performance information. There is also the possibility that the changes were already explained in the budget documents, so there was less need to explain them in the SOI and annual report.

## **Output Costs**

There are seven items, organised into two groups, that relate to the disclosure of output costs. The first group that contains information at output class level consists of four items: total actual cost, total planned cost, disclosure of cost allocation method and disclosure on variance (if applicable). The second group that lists cost information at output level consists of three items: actual cost, planned cost and disclosure of variance (if applicable).

At output class level, only quantitative information was disclosed comprehensively. Almost all total and planned cost information at this level was reported. This result was expected, since this information was easily available through the budget system, so it was simply a matter of pulling out the mandatory information from the budget process and disclosing it in the SOI and annual report. However, when other information was not available in the budget system, almost no effort appears to have been made by the agencies to disclose it in performance reporting. One such piece of information was, as discussed above, disclosure of cost allocation method. Other mandatory information, such as disclosure of variance, was also not reported in the



majority of cases. The disclosure was presented in 19% of cases for 2003/2004 to 42% of cases for 2006/2007, with an average of 35% in other periods of observation.

The comprehensiveness of cost information reported is more disappointing at output level. In all periods, less than 20% of the cases, agencies reported the information of budgeted/planned cost for their outputs. Reporting of information on actual cost and cost variance was even less. In cases of non-reporting, there is the possibility that the cost information at output level was not available in the management system of the agencies. This may also indicate that cost management was scarcely implemented at output level by most agencies. Furthermore, this information was not mandatory, thus there was less pressure to provide this type of information.

### **Quantity Aspect of Performance**

Information on indicators for quantity was disclosed quite comprehensively. This aspect of performance was assessed by identifying the presence of information on the planned quantity and the actual quantity of outputs produced and the explanation disclosed. The scoring for these items was assessed at output level. Agencies consistently reported the quantity of their outputs if the outputs were identified in their reports. Giving the scoring system used, only in a very few cases was this information not available. This could be seen by comparing the scores in the disclosure items titled “a list of detailed outputs” and “actual quantity”. In these cases, agencies did not disclose deliverables or milestones of their policy advice outputs and thus no score was given on this information.

The level of comprehensiveness found in the information presented on planned quantity and variances was not as high as that for actual quantity information. On average, during the periods of observation, planned quantity information was disclosed in only 71% of cases. The corresponding figure for variance disclosure was much lower, around 53% of cases. The low scores in the planned quantity information category may be due to the fact that this information was not available in the budgeting system. Outlines of detailed outputs to be produced are the responsibility of the agencies and their respective ministers as they prepare the Output Plans. Since this documentation is available to the public, there is a possibility that only selected or summary information is disclosed in an SOI and annual report. If someone wants to have this information in detail, they can refer to the Output Plans. When planning information is missing from the annual reports, readers may not easily assess the performance of agencies regarding their achievements in delivering the output promised at the beginning of the fiscal year.

Readers have to spend more effort on accessing and studying Output Plans in order to make an assessment. The low scores in the variance disclosure item reflect lack of consistency of reporting practice. While some agencies explained any variances reported in the documents assessed, other agencies disclosed nothing relating to the same information. It seems to be a matter for each individual agency's policy as to whether or not that information is reported.

### **Quality Aspect of Performance**

Eight items in the disclosure index, which are arranged in three groups, represent the quality areas of performance. TPA-9 requires only two out of these three groups of items: quality description and quality measures. Information on quality benchmarks is included in the index because TPA-9 shows a preference for quality measures. Measures that are externally benchmarked and assessed are the best measures. Thus, if an agency used external measures for the quality aspect of performance, this agency was considered to disclose better information. It was also noted during the development of the index that if agencies were reporting externally benchmarked quality measures, they also reported internally benchmarked measures. These are the reasons for incorporating three items in the group of information on the quality benchmarks used, so that the index can produce a higher differentiation in assessments.

The level of comprehensiveness of information presented on quality areas was still quite high, with average scores ranging from 60% to 92% (refer to Table 2 of Appendix E for the detail), but it was lower than the level of comprehensiveness of information disclosed on output description, cost measures and quality measures. Agencies disclosed more comprehensively on the list of quality dimensions, but they disclosed less when describing dimensions and presenting measures. There was a high degree of preference for using external benchmarks for quality measures. In four out of five cases of reporting quality measures, an external benchmark was used. When this benchmark was used, it was usually assessed externally, mostly by customers of the agencies.

The scoring results in this area of performance reporting showed that agencies were concerned with the quality of most of their outputs. Planned and actual measures were reported in more than 75% of the cases. In addition, when there was a variance between the planned and the actual indicators, agencies tended to give reasons for any

variance. Even a very small inconsistency was reported. In a small number of cases, either the targeted measures or the actual measures of quality were not reported.

### **Timeliness Aspect of Performance**

The level of comprehensiveness seen in this area of performance reporting was lower compared with other areas, except for the reporting of cost information at output level. This aspect of performance reporting was assessed by three items in the index: the planned and actual measures of timeliness and the explanation of variance. There was a small number of inconsistencies found when these measures were reported on; sometimes planned measures were not reported at all in cases that did report actual measures. Similar to previous cases of reporting inconsistency, this inconsistency may stem from the unavailability of relevant information for the planning process. It may be difficult to set a timeliness target in advance for some activities/outputs undertaken in the public sector.

## **2. Outcome information**

The level of the quality of outcome information disclosed, as presented by scores shown in Table 3 of Appendix E, was very low. After 5 years of the MfO initiative roll-out, on average the amount of outcome disclosed was only around one-third of the complete set of outcome information prescribed by guidance documents. The information quality also changed very little from the first accountability documents reported under the MfO initiative to the last period analysed in this study. Based on average scores, there was a noticeable rise in information disclosed in the first year, and then a small increase was noticed in the second year, but the quality then remained the same for some periods, and lastly there was a small increase in the fifth year.

The pattern of quality improvement over time was seen mostly in very small gradual increases, except for some agencies. Five agencies that scored the highest in the first year of observation were agencies that perform a great deal of service delivery, except for Treasury. Three agencies, Ministry of Health, Ministry of Social Development and Ministry of Research, Science and Technology, exhibited a noticeable jump in the comprehensiveness of information reported during the second year of observation. This may indicate that these agencies learned from their first year of experience to do better in the second iteration. The Department of Labour's scores decreased over the period of observation. The Ministry of Health and Ministry of Science, Research and Technology showed scores that increased during earlier years

and decreased in later years. Archives New Zealand was able to substantially increase its information comprehensiveness in the last year of observation.

The scoring results per disclosure item presented in Table 4 of Appendix E reveals that agencies were strong when reporting qualitative information. This information was also found to have been developed during the earlier stages of the MfO initiative. In the first year of the period it was clear that agencies were able to define the outcomes that they were trying to achieve. They were also able to further elaborate their outcomes and presented them in hierarchical order, depicting low-level (short-term) outcomes and high-level (long-term) outcomes. In most cases, they also explained their organisation's objectives/goals/purposes and linked them with both their organisation's outcomes and its health and capability. There was an indication that what agencies did at the basic level was to take the reporting format provided, fill it with simple narratives that were easy to develop, and add a more difficult narrative and some quantitative data, if they were able to provide that, in later years.

No agency presented comprehensive information on intervention. Almost all agencies described the link between objectives and outcomes and between low-level outcomes and high-level ones in a very simple way, either as a short narrative, or in the form of a tabular or graphical presentation. There were a few attempts to elaborate the links in the early years of the period, but later these attempts were abandoned. This seems to indicate that even at outcome level agencies were struggling to understand the relationship between short-term outcomes and long-term outcomes. A similar finding was also seen in the disclosure of intervention logic of outputs related to the intended outcomes. Most agencies showed intervention logic in simple assertions, either in narrative, tabular or graphical presentations. There were a few agencies who attempted to elaborate on why and how the outputs would contribute to the achievement of outcomes. However, the elaboration did not go far enough to show that these agencies had completed sufficient analysis of various potential outputs to see how these might achieve the targeted outcomes.

Many guidance documents and tools available stipulate that the purpose of performance reporting is to tell a performance story. The tendency of agencies was to follow only the reporting format outline in the guidance documents, instead of developing and reporting outcome information based on the idea of telling a performance story. There was a lack of disclosure of the various output alternatives considered, when selecting the best mix of outputs to achieve the targeted outcomes. Discussion of factors affecting outcomes was also missing in a large number of the

documents analysed, which may imply that outcome-based management had not been properly implemented, so as to enable agencies to come up with a comprehensive performance story. Around a third of the agencies identified these factors in later years of observation, but very few of them provided any further explanation of these factors. Even worse, none of them tried to explain their projections in a way that indicated that they had anticipated these factors when pursuing their outcomes.

Information about the cost effectiveness of interventions and its indicators, which was mandatory after the PFA was amended in 2005, was poorly provided. In the last two years of observation, more and more agencies were starting to give some information on cost effectiveness, but in the last year of observation the number of agencies reporting this information was still less than half. When presenting information, few provided quantitative information.

The provision of quantitative information was most deficient in the area of outcome information. While output information was dominated by quantitative measures, outcome information infrequently contained such measures. The few agencies that disclosed measures, mostly presented information from previous years. Measures of cost-effectiveness and capability were almost non-existent. Planned/targeted measures were not found in the accountability documents. This may highlight two possible factors: first, the agencies' information systems were not developed sufficiently to enable them to produce this information, or second, the outcomes were too complex to be measured.

A detailed analysis of the score results for disclosure items of outcome information is presented below. The discussion is divided into five parts: goals/objectives identification, outcome description, outcome target/indicator descriptions, intervention logic discussion, and presentation of indicators.

### **Goals/Objectives Identification**

In the Public Finance Act, there is a distinction made between goals/objectives and impacts/outcomes. The Act recognises that not all outputs or activities or functions of a department, such as ministerial servicing, regulation drafting or departmental chief executive's appointment, are intended for pursuit of outcomes. These outputs/functions are not intended to produce direct societal, economic or environmental impact. Therefore, some agencies may include objectives in addition to outcomes to provide alternative targets.

In practice, there was some degree of confusion among agencies. Instead of listing objectives alongside outcomes and linking them to some outputs, these agencies disclosed goals/objectives separately as the highest level of targets, then they listed and explained the outcomes that they pursued as intermediate targets to achieve their goals and objectives. To take into account this variation, in this study the links identified were not only links made between goals/objectives and outputs, but also between goals/objectives and outcomes. Therefore, the scoring results in this section should be interpreted with care.

Overall, the comprehensiveness of the information reported on goals/objectives was quite high. On average, in 75% of cases, agencies disclosed their goals/objectives. The score cannot be interpreted to show the cases in which there was no identification of outcomes of the agencies' activities/outputs for the reason mentioned in the previous paragraph. It can only be interpreted as the percentage of cases in which the goals/objectives of the agencies were identified. It cannot also be interpreted that in 25% of cases, there was no identification of the goals/objectives of the agencies

An interesting observation is made concerning the information revealed about goal/objective linkages. In more than 70% of the cases during the period of observation, agencies that disclosed this information linked their goals/objectives with outcomes instead of outputs. This practice was quite confusing, given that the idea of the inclusion of goal/objective identification was to provide an alternative to outcomes as a target for output delivery. The percentage of cases in which agencies reported this link of goal/objective and outcomes was quite high and showed an increasing trend, from around 72% on average for the first two years of the period to 90% in the last year. In addition, there seems to be some degree of difficulty faced by agencies in disclosing the linkages. All of those that disclosed simply asserted the link without giving an elaborate explanation of why there was a link. The same finding is also noted in the disclosure on output-outcome linkage, which is discussed below at page 129. As previously stated, the OAG also found the same problem with the practice of linkage disclosure when reviewing the performance reporting practice of the agencies (Controller and Auditor General, 2008).

### **Outcome Description**

In terms of the comprehensiveness of information presented, the highest quality of information was found in the disclosure of outcome description, except for the linkage discussion component. There was an increasing trend to report this information

more comprehensively over time. In all years of observation, only one agency did not disclose its outcomes and that was only in the first year of the period. Most of the outcomes disclosed were explained. The trend for disclosing an elaborate explanation of outcomes was an increase from around 75% in the first year of the period to 96% in the last two years. In the other 4% of the cases each agency seemed to conclude that, a short description was enough to properly inform their readers.

The ability to produce an outcome framework, which outlined high-level and low-level outcomes in a graphical format, was seen as evidence that the MfO initiative was going in the right direction. Since most agencies wanted to be seen as making progress in their efforts to apply outcome-focused management, it was not surprising that there was quite a large number of cases in which an outcome framework was presented in the documents analysed. It seems that some agencies were facing difficulties in developing the framework in the first year of the MfO initiative, in which the framework was reported in only 59% of cases. In the next four years of observation, this figure increased to around 80%. The lack of an outcome framework in 20% of cases may be associated with difficulties in developing the framework or to the fact that the outcome was simple, so there was no need to elaborate it further by describing various levels of outcomes. However, similar to other parts of disclosure in which the links needed to be explained further, there were few attempts by agencies to provide more explanations. These limited efforts were seen only in the first three years of the MfO initiative.

### **Description of Outcome Targets/Indicators**

The scoring results for this section showed that ministries/departments were struggling during the learning process to come up with outcome targets/indicators. There were many more attempts to present targets qualitatively rather than quantitatively. In the first year of the MfO initiative, in only half of reported outcomes were the targets qualitatively discussed. In three out of four cases, the targets were multi-dimensional. In the fifth year of the MfO initiative, there were 82% of cases in which the qualitative targets were reported and 80% of them were multi-dimensional. Thus, more and more agencies were able to set outcome targets over time, indicating that there was a learning process going on during the implementation of the MfO initiative. The limited findings of the early review of the Pathfinder Project also indicated the difficulty. In their review, the Economic and Strategic Group (2003) stated that six agencies involved with the Project, which was supposed to help them develop

outcome targets, could still not point to specific (outcome-related) results in their activities just one year before the MfO initiative was mandatory for all agencies.

There was an indication that performance in the outcome area was not observable, making it difficult to set targets, even qualitative ones. There were still many cases of reported outcomes in which no targets were disclosed. The level of disclosure for quantitative targets for outcomes was very low. This information was not disclosed in more than 21% of cases in all years of the period. The difficulty was also noted by the Economic and Strategic Group (2003) in their review. The six agencies involved in the Pathfinder Project, which was a laboratory project for the MfO initiative, still could not come up with performance measures for outcomes. The review revealed that these agencies needed to assign significant effort or resources in order to be able to come up with outcome measures in their management systems, so that they could report them in the corporate documents.

### **Intervention Logic, Output Alternatives and Other Factors Affecting Outcomes**

Intervention logic is defined as ‘a systematic and reasoned description of the causal links between a department’s activities, outputs, immediate and end outcomes’ (The Pathfinder Project, 2003b, p. 1). Proper intervention logic contains at least two components: the outcomes to be achieved and the logic and evidence of the link between outcomes/goals and outputs. Intervention logic that is well developed shows the justification for the choice of outputs and the improvements in outcomes based on the effectiveness of the outputs selected. In the guidance document produced by the Pathfinder Project, there is a list of management applications of intervention logic. These applications, among other things, include the prioritisation and alignment of outputs to maximise outcomes and the identification of risks (factors) affecting outcomes or goals. When considering its definition and application, it seems that a comprehensive explanation of intervention logic would consist of the assertion of the link, the explanation justifying the link, discussion of outputs alternatives to show the prioritisation and alignment of outputs, and discussion of risks or factors affecting outputs and outcomes.

A compliance motive may be the overriding purpose driving agencies when reporting performance information in their corporate documents, instead of the priority of providing readers with the necessary information relevant to their assessment of an agencies’ performance. If the purpose of reporting outcomes is to provide information on how an agency is effectively managing its outputs to achieve its desired outcomes,



the report should tell a credible story of performance. There should be an explanation of why the current outputs are selected instead of other alternatives. Since outcomes are less controllable by an agency, there should be some information on what other external factors influenced outcomes, so that readers of the report can assess the level of the agency's ability to control outcomes. If an agency's motive is to comply with reporting requirements, what it tends to do is disclose the minimum level of information required by regulations or guidance documents. In relation to intervention logic, central agencies' guidance on the preparation of an SOI requires agencies to discuss the links between outcomes and outputs and the risks involved in managing the delivery of outputs. If compliance is likely to be the main motive governing the preparation of agencies' reports, it is likely that a discussion of output alternatives and factors would not be widely available in the documents analysed.

The scoring results show that two factors may influence the level of disclosure in the area of intervention logic information. First, there was strong evidence that compliance was the main motive. Agencies tried to report all mandatory information for the intervention logic disclosure. In an average of around 88% of cases during all years of the period, agencies disclosed the link between outputs and outcomes. However, no attempt was observed to disclose non-mandatory items of information, which in this situation meant a discussion of output alternatives. Second, there seemed to be some degree of difficulty faced by agencies in reporting the information. Not much discussion was presented to justify any links. Most of the presentations of intervention logic were in simple form, such as a simple assertion or a tabular or graphical presentation. Furthermore, the discussion of factors influencing outcomes was very limited. At best, agencies identified the factors in less than a third of cases, even though the trend indicated an increase from 26% of cases in the first year of the period to 37% in the last year of the period. Discussion of these factors in agencies' reports was infrequent. No discussion was presented on management's assessment of these factors. All these findings may indicate that the reporting of intervention logic is one of the most difficult parts of the MfO initiative to be developed and implemented. The findings may also indicate that there are some obstacles in the process hindering a proper implementation of managing for outcomes in New Zealand government agencies.

The review of the Economics and Strategy Group (2003) showed difficulties with the development of intervention logic. In the early stage of the MfO initiative, this Group reviewed all of the SOIs produced and concluded that the alignment of business

plans to outcomes in the SOI was recognised as a requirement, rather than as a reality, in the operations of government agencies. There was a standardised broad structure of intervention logic identified across agencies. Attempts to show the alignment, as acknowledged by senior managers when they were interviewed by the Group, were merely in the form of mapping existing votes and output classes to the new outcome structure. There was an expectation that the next stage of alignment would occur, in which the determination of a combination of outputs that better supported the priorities would be reflected in the outcome framework. However, this expectation never materialised.

### **Presentation of Indicators on Outcomes, Cost Effectiveness of Intervention and Organisation Health and Capability**

The PFA requires an extensive list of indicators to be disclosed in annual reports for outcome reporting. There are three groups of indicators: outcome indicators, cost effectiveness indicators and organisation health and capability indicators. At least two types of indicators should be reported for each group: the targeted indicators or standards and the actual or result indicators. One important issue that needs to be considered in reporting outcome indicators is the need for comparisons, either historical comparison or inter-organisation or inter-jurisdiction comparison (The Pathfinder Project, 2003a). Therefore, a complete set of outcome indicators would consist of targeted indicators, actual indicators in the current year, indicators from previous years and indicators from other organisations/jurisdictions. There should be an adequate explanation of these indicators. For indicators of cost effectiveness and organisation health and capability, the requirement for the set of indicators is not clearly spelled out. The only requirement, as stated in the PFA, is the disclosure of the main measures and standards.

The scoring results of these indicators clearly showed that it was not easy for public sector agencies to come up with quantitative information for outcome reporting. Outcome indicators were reported in less than 20% of cases during the years of the period of observation, except for the second year, in which the figure was slightly higher than 20%. Agencies may not have information systems which support the disclosure of these indicators. There was an indication that if the information was available in their systems, agencies would report it in their corporate documents. If agencies were able to report the current indicators of outcomes, they would be likely to report the comparative indicators in previous years. The pattern in the scoring results

may indicate the main factor for reporting of the indicators. There was no evidence of any outward trend in the disclosure of these indicators. This may indicate that agencies were struggling to come up with a quantitative measure of outcomes.

The disclosure on cost-effectiveness indicators was most disappointing. Even though it is required by the Act, the indicators were barely reported on in the corporate documents issued by the agencies. Cost-effectiveness indicators describe the link between outcome indicators and resource/effort indicators. Therefore, agencies needed to be able to produce outcome indicators first, before they could report cost-effectiveness indicators. Since the level of information reported on outcome indicators was very low, so was the level of information reported on cost-effectiveness indicators. Some agencies came with measures for cost-effectiveness by disclosing descriptions of how cost-effectiveness would be measured, by using qualitative indicators. Furthermore, there was little elaboration on why these indicators were selected. On the other hand, there was a promising trend of more and more agencies reporting cost-effectiveness indicators. Cost-effectiveness measures were described in only 4% of cases in the first two years of the period. That figure was 40% in the last year of the period. Still, it was a disappointing result, since most of these agencies were not able to properly disclose their cost-effectiveness even qualitatively.

Slightly better results were found in the reporting of organisation health and capability indicators. Agencies reported these indicators qualitatively in more than 67% of cases in the first year of the period. This figure increased to around 90% in the last year of the period of observation. Thus, there was an increasing trend of more and more agencies proposing the indicators needed to measure the health and capability of their organisation to deliver outcomes. Unfortunately, this was the most that these agencies provided. There was not much discussion on the proposed indicators. Furthermore, no quantitative indicators were reported.

## **B. Findings from Interviews**

The findings from the interviews are presented in several groups that represent the main themes identified during the analysis of the transcribed record of the interviews. When available, results from other published reviews are also presented for a comparison with the interview findings.

## **1. MfO Initiative, Its Intention and Impact on Performance Information**

According to the interviews with officials from the central agencies, the MfO initiative was introduced following a series of initiatives undertaken by central agencies to support government agencies to adopt a result-oriented management system in their policy advice and service delivery operation. The MfO initiative was preceded by two pilot projects. The first one was the Capability, Accountability and Performance (CAP) Pilot and the other was the Pathfinder Project. The interviewees explained that these two pilot projects, especially the Pathfinder Project, were bottom-up processes in which participants led the projects and the central agencies supported the process by setting up a secretariat to facilitate the coordination of activities. Participants chaired all processes, both at the steering group level and working group level. When senior leadership in the government saw that through the Pathfinder Project an outcome-oriented management would be workable, central agencies decided to run the MfO initiative in parallel. Thus, MfO is a top-bottom initiative, under the direction of central agencies, initiated to apply the knowledge learnt during the Pathfinder Project to public sector agencies.

The first official interviewed argued that the main idea of the MfO initiative was to change the management approach in public sector agencies. The production of outcome information was viewed as a by-product of the process which gave an indication that the result-oriented management approach was being successfully applied. It was not about accountability for outcomes but it was about the responsibility of chief executives to help ministers understand the outcomes being achieved by their agencies. In an interview, Respondent 1 from central agencies stated that:

MfO is a managerial process ... [and] involves ministers as well. So, it's not just departments and their silos but it's about open discussion, a willingness to change, a good technical appreciation of how you deliver results within your environment. But, it's not a spotlight from the press, shining on what is a problem and so on. (Respondent 1)

The MfO initiative reflected a change from the public sector management previously established during the first public sector reform of the late 1980s that put the onus of accountability for outcomes on ministers while accountability for outputs was placed on chief executives. The arrangement whereby ministers were accountable for outcomes did not work in practice, as ministers did not have systems to identify and monitor outcomes. Respondent 1 described the situation as follows:

The only people that can help ... the minister are really the departments. And if they fail in their duty, then the ministers cannot really be held responsible for things that [they] are accountable for ..., as [they] need

information about outcomes. [Ministers] need information about the way in which outputs link to outcomes. And without that, [they have been given] a job without giving them the means to perform. So, what the difference in the MfO [means was that departments] retain managerial accountability for output with obviously ministers signing up to output agreement and so on, but the departmental managers themselves, the chief executives, and so on, and through their reporting obligations, the entire department reporting to parliament becomes accountable for not so much the delivery of outcomes but they become accountable for the processes that should lead to outcomes. (Respondent 1)

With the MfO initiative, a new set of performance information focusing on outcome achievement was being reported, not for the purpose of making chief executives accountable for outcomes, but for showing that outcomes were being considered in the selection of outputs to be produced, as this interviewee stated:

...the chief executive's accountability for outcomes ([s] along the lines of know what your outcomes are, ... measure the outcomes, have a ... sensible way [of] linking outputs(particularly) the big things you spend money on, to those outcomes, to collect information on that and to respond sensibly to the performance information you get. (Respondent 1)

Several things were expected to result from the MfO initiative. The first interviewee argued that there would be more information coming into the system to support the perpetual debate between central agencies and auditors on one side, and the departments on the other side, on how to allocate resources properly to get the results intended. He further stated that in the MfO environment, a coupling between financial and non-financial information could be achieved, to allow for sensible decision-making in resource allocation within a public agency and across agencies in the public sector. Second, the MfO environment would foster evaluation processes in an agency, to understand what is working and what is not working and to use this understanding to propose a different direction for the organisation's strategy. Third, the MfO would encourage agencies to work together, as they realised that their outcomes were shared by other agencies. This, in turn, would reduce the dysfunctional silo effect resulting from output-based accountability.

## **2. Stakeholders' Interest in Performance Information**

The response from stakeholders to the performance information reported in the accountability documents was very small. All interviewees stated that auditors were the only ones that consistently pressured them in terms of providing better information over

time. Ministers demanded some performance information, but in the majority of cases it was not that which was reported in the accountability documents. It seemed that members of parliament never give performance information any attention, since it was not likely that they would find anything that could be used to score points in political debate. Central agencies insisted only that agencies should show improvement in the quality of information produced over time, without any specific milestones set in advance. Any other stakeholders, such as users of the service or society, were not mentioned as having an interest in performance information.

Ministers' attention to and expectations about reported information were very low. However, in some cases, they showed an interest and were willing to be more involved in the reporting process. As predicted in agency theory, if there is a possibility that any disclosed information might be considered bad news or a negative result or invite a negative public response, then ministers would get involved in the reporting process more frequently. One top-level official stated that:

[When] ministers get involved frequently that is the case that points us to where it does not work. Where the interventions or the measurements show ... negative trends ..., that is where the ministers get a sense ... that [a] public response might come out of that. (Respondent from Agency I)

Another area of information that ministers were generally interested in is output information. This information relates to the output plan that they and the chief executives have agreed on at the beginning of the fiscal year. With some agencies that produced policy advice outputs, there was a strong indication that ministers were trying to exert their control as purchasers by demanding information about what they could get for the allocated budget. This is quite rational for ministers, as they are the purchasers and users of the outputs provided. As purchasers and users, ministers want their concerns and priorities to be addressed by their agencies. Scrutinising the output information is an important way of ensuring these things are properly delivered. One interviewee from these types of agencies summarised this interest by saying:

They want to see the things that are their priorities for the year within the output report. They want to know that they are paying the [right] price [for what] they get. (Respondent from Agency H)

When agencies requested a significant increase in the budget or funding, ministers would ask for more information to justify their support for the budget increase. They were interested in outcome information, interventions proposed and the cost of intervention relating to any increased spending. They wanted to dig deeper to get

more information and they demanded that the information be clearly described, not only in any funding proposal circulated to the select committee, but also in the required accountability documents, especially those issued during the year the increased funding was granted. One top-level official in such an agency explained it this way:

...in 2002, we were preparing a business case for a reasonably significant increase in funding to the Ministry. ... So... when we were preparing our annual performance information and financial information, there was a strong focus on setting out what ... we [were] wanting to achieve for the next five years, i.e. ... our strategic priorities, and how much would that cost. So there was quite an engagement with the Minister of that time. (Respondent from Agency H)

Lastly, ministers of small agencies showed their interest in outcome information, for the purpose of defending the existence of their organisation. The existence of these small agencies continues to be at political risk, either of their being incorporated into another agency or being dismantled. Their ministers would try to ensure that the planning documents showed their priorities so that they could support the existence of their agencies, because these agencies were pursuing outcomes that were important for maintaining or gaining the political support of certain constituents.

One possible explanation of why ministers' interest in the accountability documents was low is that they could access chief executives easily for information. As one respondent put it:

... in day to day things ..., the chief executive and the minister will agree [on the] things we want to do, where we want to go, and so on, [and that was] perhaps more important than these documents. (Respondent from Agency E)

Another possible reason for low interest in the performance information is that management control systems still concentrated on the most controllable aspects of the system, which was input control. As one interviewee from the central agencies put it:

And ministers never allocate resources on the basis of outputs. It's always on the basis of inputs, so salary costs go up to pay for that, or we need a new IT system, [so] we pay for that, and I need more staff, [so we] pay for that. But, they never sit down and say, "Look we need more of that and less of that." And it is up to the CEO to decide how best to produce that. So, although we [have] got this outcome output focus in the legislation, on [a] day to day basis, almost all budget decisions are made on inputs not outputs. (Respondent 3)

Parliament, or in this case members of select committees, showed little or no interest in performance information. The ones who did show interest were members of the Opposition wishing to dig up bad news, as one top-level official put it:

...question by Opposition but more on the retrospective side, more on negative results on the reported information... (Respondent from Agency A)

Another top-level official elaborated on this tendency:

...a select committee is not a part of [the] political setup [conducive] to openness in the government .... in the end it is [simply] more political games ... without real benefit to us. [It] does not challenge us to be able to develop ideas. It challenges because it is trying to score political points. (Respondent from Agency I)

Central agencies were cited as the main actor for demanding improvements in performance information. The SSC concentrated more on getting chief executives to pay attention to outcomes, while Treasury focused on the outcome linkage of price to quantity or the cost-effectiveness of an executive agency's intervention. Treasury particularly had more leverage in demanding information through the budget process, as output information went into government estimates and appropriations. One interviewee, who was involved in some budget development processes, summarised the situation:

Treasury [was] ... too focused on price, quantity and efficiency but [took] the lead in outcome linkage. Treasury tried to use leverage in terms of budget allocation, but ... in subtle ways, [it looked] at the capability and performance information as indications for allowing some agencies to grow in budget, especially the ones [whose goals were] being pursued by the government. (Respondent from Agency A)

Another interviewee argued that central agencies should engage in more explicit action to get all elements of the New Zealand public sector management system to focus on outcomes. The system should provide strong incentives for agencies to improve the quality of outcome information. The audit function has done that to some extent. However central agencies should provide further pressure. He elaborated on his opinion as follows:

So, if the outcomes are the central organising [principle] or focus for public servants, who is basically reviewing the quality of them? ... I do not think central agencies are .... I do not think they [have] got the capability. Is it a function they should have? Yes. (Respondent from Agency G)



Auditors maintained their interest in performance information because of their mandated role of auditing the Statement of Service Performance (output information). They have an obligation to audit output information, so obviously there were a lot of discussions between agencies and their auditors regarding the presentation of output information. In auditing this information, auditors generally approached their job by concentrating on ensuring that the reported output information was reliable, in that it was corroborated and produced by a sound management system. They took a lighter approach when considering the appropriateness of output information in the accountability document. They stated in an interview that Treasury and the agencies themselves were the responsible parties for determining the relevance of the information, since they incorporated this information into the budget process.

However, auditors started to show an interest in other parts of performance information and in recent years they criticised individual agency for the quality of information provided, as indicated in their management letter sent to the chief executive, minister and select committee of each respective agency. However, this assessment was not publicly available. The one assessment that was publicly available was a general assessment from the OAG made about the quality of performance information as seen across the public sector. This assessment was published as a Discussion Paper. However, the paper did not contain information on the quality of performance reporting for individual agencies. This shift in the auditors' interest in performance information was echoed in the statement of one interviewee who said:

Audit had criticised and issued reports and now is starting to mark department's documentation and grade it... They criticised, that's useful, but they cannot tell us what to do. We have had to work it out for ourselves. (Respondent from Agency B)

The consistent pressure of the auditors by means of auditing and reviewing performance information resulted in a strong incentive for government agencies to follow auditors' preferences on the types of information to be presented and the way it should be reported. One top-level official described this tendency this way:

So, in the management letter they give [a] grading. That management letter goes to the minister and goes to [the] select committee. When [the] select committee looks at the annual report, there is a very strong incentive to put something in the place to keep the auditor happy as opposed to [whether the information] is good quality or not. (Respondent from Agency I)

As auditors focused more on reliability over relevance of information during the period covered in this study, it had the potential to create a situation where the information provided might not be that which is appropriate for the intended audience of accountability documents. It could also contribute to the presence of cross-sectional inconsistency in the reporting.

The review by the Economics and Strategy Group (2003) of the first iteration of the SOI preparation under the MfO initiative showed some dialogue and engagement by the agencies with their stakeholders. The initial discussions on outcomes were mostly held internally at senior management level. There was not much engagement with staff at levels lower than the second tier of management. External engagement was usually started once chief executives were comfortable with their outcomes statements. The level of this engagement varied from one agency to another. Ministers were effectively involved with output planning, but they were reported to have limited engagement in the preparation of outcome performance information for the SOI. Two obstacles were cited as the reason why the engagement of ministers was very low. First, there was a concern that some ministers did not have adequate knowledge of, aptitude for and/or interest in the planning process of their agencies. Second, some agencies faced difficulties in engaging with multiple ministers with quite different policy agendas and operational priorities. It was logistically difficult to get all ministers together to discuss relevant issues at the same time. The process of separate consultation could intensify fragmentation. Therefore, some agencies considered that the best way to develop an outcome framework was to do it in-house first, with ministers' engagement later for review and consideration purposes.

### **3. Performance Information and Organisation Achievement**

All respondents were of the view that accountability documents are an important means of communicating an organisation's performance to stakeholders and to gain public trust and confidence in an organisation. However, they believed that the Statement of Intent and the annual report were not adequate if they were being used to assess the organisation's performance. Readers also needed to look at an Estimates and Output Plan in order to understand the overall picture of what an organisation had done. Nevertheless, the respondents contended that their accountability documents described the major achievements of their organisation. This assertion was stronger concerning output information relating to service delivery. However, for other non-financial performance information, they offered far less assurance. This level of assurance was

related to the notion of the organisation's ability to control the results. Therefore, they tended to be less explicit in spelling out the outcome they were trying to achieve and whether or not they had achieved it. One interviewee described it by saying:

... I think one of the fundamental issues with the public management system that came in because of the PFA is that chief executives necessarily, in my view, move to define what they can manage, what they [can] measure, what they can deliver, [and] what they can guarantee. So... in an environment of uncertainty, you become very limited in your vision. (Respondent from Agency A)

In terms of policy advice, the quarterly monitoring report on the output agreement, which was available only to ministers, better reflected the organisation's performance in delivering policy advice outputs. Measures reported in the accountability documents were usually a few selected measures that were made available in the output agreement to satisfy accountability requirements. Business decisions were not based on these measures, and as one interviewee put it:

...service performance measures are not driving our output decisions, so business decisions [based on] what you are doing are not really relevant for service performance measures. (Respondent from Agency G)

The output agreement and the monitoring report provided clear targets of performance and showed the organisation's contributions, so that people inside the organisation could draw a link between activities and outputs. One interviewee explained it this way:

Our main measure of have we done what we said is actually our reporting against the output agreement with the Minister, which for us is done now quarterly and monitored quarterly by senior management team ... the business is our output agreement [with] the Minister and we're actually having more detailed [input into] that ... we actually have our group business plans, which [have all] that we're doing in the Ministry and we monitor ourselves and then report against that. (Respondent from Agency F)

#### **4. Development of Output Measures**

Agencies developed all indicators for output reporting by themselves, with minimal involvement from stakeholders. Ministers were involved mostly in setting direction and sometimes signing off reports, especially for the planning documents. Auditors maintained an interest in the reliability of the reported information. Central agencies were involved indirectly by issuing guidance documents to be used as a reference when reporting. Professional guidance (TPA-9) was not mentioned as being used in the reporting process.

Ministers were involved at the early stage of the planning process by setting priorities. They set goals and selected some areas to focus on for the next fiscal year. Chief executives prepared all the detail in the corporate plan, including the indicators to be used and the targets set for each indicator. Ministers appeared to accept most of the agency's proposals. The typical process was described in an interview as follows:

I would say it is not a structured process... what we do is the Minister sets the goal for the year. ... He [sets] a limited number of priorities..., [he gives] agreement on some plan for doing it but it might not [be a fully detailed] plan, fully elaborated and ... the job of the Ministry is to work out how to get there. So, at the operational level, we choose the outputs. Now, some of [that] would be opportunistic, sometimes you set something [to happen] that you couldn't foresee when the planning was done. (Respondent from Agency B)

Thus, the selection of output measures was very much at the discretion of agencies, with no influence or control from ministers. This process raises a concern that agencies might select measures that would not show bad results, rather than select measures that were appropriate to their activities. As one top-level official put it:

I guess [this involves] quite a bit of trust in the reporting entity that they picked the ... right ones and they are not hiding something. They are not avoiding [a] quantity measure that shows they are not doing well. (Respondent from Agency I)

Auditors also did not play any role in ensuring that the measures adopted were appropriate. They audited what was reported and provided an opinion on the reliability of information presented, but not so much on the relevance of the information provided. The auditors' approach was outlined in one interview:

... OAG and through their auditors took the approach that basically what you come up with, whatever you [are] going [to] come up with ... we will provide an opinion on it. (Respondent from Agency I)

An interview with officials of the audit function revealed that:

We've attested that they [performance measures] fairly reflect the forecast set .... And that was on the basis that we assume that others, such as Treasury and SSC, were helping [agencies] to think about what is appropriate. ... And, it's [at the] discretion of preparers to think about and approach intelligently, as [Respondent 5] says, thinking about what I need [to report] to be accountable to Parliament and the public about. This is taxpayers' money; how do I demonstrate to them that I am managing that output well? (Respondent 4)

However, over time there were indications that auditors were becoming more interested in the architecture of measures and standards. This pressured some agencies to do more

in some cases, even hiring consultants to help them come up with better measures for outputs.

Four respondents from two agencies argued that since output information development was part of the budget process, there was a robust process already in place to ensure that the measures selected were fit for the accountability purposes of the government budget. One top-level official argued for the appropriateness of selected measures by saying:

I think [with] outputs, the coherence of output structure is really strongly robust because it is driven [by] the framework for financial management. It is [a] reasonably robust legacy of thinking and technical expertise ... [it is] typically properly managed through [the] budget and costing mechanism..., which aggregate outputs. (Respondent from Agency A)

Since output measures are developed as part of the output plan and output agreement development, measures are available in the information system to support the budgeting process and subsequent accountability reporting requirements. On the other hand, the information would be high quality only at output class level, the level required by the budget process. There was no strong pressure for the development of more refined output information, such as availability of measures at output level, with cost and an organisation's resistance being cited as the main causes for this absence. While the current information system can provide some information on quantity, quality and timeliness indicators, there was not much information on the cost of each output, which required agencies to set up a sophisticated costing system. In an interview, one top-level official revealed this concern:

We talked internally just in the last few months about the need to get a much greater level of granularity around what outputs we produced below the output class level ..., the information at the moment [is] unsatisfactory and it comes with [a] compliance cost and to be blunt about [it] some organisations resist ... the management system that [it needs]. (Respondent from Agency A)

Another interviewee even acknowledged that the measures that were reported in accountability documents were not useful internally. He stated that:

One thing we found at [Agency C], if I go back to that as an example, was that early on we had some output measures which were being measured and reported [on] only in the annual report. So, therefore it was a nuisance [to have these output measures and required], an additional compliance effort from the people around the country. So, they did not care. They were not capturing good quality ..., robust data. They would just [comply] because they had to do it because that

[information went] into the annual report, and they were not using that, [as it] was not useful to them. (Respondent from Agency C)

In general, selection of output measures focused on determining the processes by which agencies collected the data to be able to justify and prove that they had the systems. If agencies did not put a priority on upgrading their information systems, that would be reflected in their accountability documents.

When respondents were asked about the reporting standards required as the reference for measure selection and output reporting, they only mentioned Treasury's guidance for preparation of the Statement of Intent and Annual Report. Only one respondent acknowledged the presence of TPA-9 which provides guidance for service performance reporting. Even respondents from central agencies did not have any knowledge of this reporting guidance. The respondents from the audit function were quite surprised at this finding, since they assumed that the accountability reports were developed with reference to TPA-9.

## **5. Development of Outcome Measures**

Agencies developed all indicators for outcome reporting by themselves, with some degree of involvement from stakeholders. As in the case of output reporting, ministers also were only involved in setting their agency's direction in advance and sometimes signing off its planning reports. Auditors assessed the reported information in later years and reported back to the management in a management letter. In some cases, central agencies were involved actively by way of facilitating the process of outcome framework development, especially in the early stages of the MfO initiative. In addition, they also issued guidance, to be used as a reporting reference. Some agencies worked together with other agencies to develop shared or sector outcomes. There was no involvement of other stakeholders, such as society and the users of services.

Similar to the process of the development of output measures, ministers were also passively involved in the development of outcome measures, even though theoretically they were responsible for outcomes. In most cases relating to policy advice on output delivery, they were involved in the process in the early planning period by providing direction on priorities and a general description of targets or objectives. It was up to agencies to frame them into outcome statements or targets. In these cases, ministers signed off what would be reported in the Statement of Intent, including outcome information. Thus, the process of outcome measure development was

essentially a top-down approach. A typical process was described by one interviewee as follows:

So this is our process: direction setting at the highest level from the government, confirming with the Ministers, developing a written understanding of what the priorities are and translating that into the Statement of Intent, and developing operational plans for twelve months so that everyone throughout the ministry, not only the units, the business units, but all staff in every unit has six key tasks which are related to the goals of the business unit. (Respondent from Agency B)

For outcomes of service delivery, the approach was mostly bottom-up, in which units within an agency came together to define what their success (outcomes) would look like in each of their areas, how each of them would contribute to the success, and what the indicators for that success were. However, the process stopped at outcome definition and descriptions of indicators; no quantitative measures were developed. Ministers also signed off what was reported in the Statement of Intent. Here is a description by one interviewee, from an agency that provided extensive service delivery, on the process of the development of outcome measures:

And, when the first SOI was established in 2002, there was a process that we went through to say what are the... outcomes we would be looking for in relation to let's say [an area]. And, then the question was how we can contribute to that, so working out the intervention logic. And then, how would we know if we're making a difference?. ... that's just a fundamental and so how would we know ... the measures [used] should [tell us]. (Respondent from Agency E)

Overall, ministers' engagement in the development and the selection of outcome measures was minimal. However, an interesting development happened recently outside the periods covered in this analysis. One interviewee revealed that in his agency the minister himself selected a set of medium-term outcomes, forced them on the system, monitored them periodically and published the measures publicly. This may be a sign that ministers are starting to engage more in outcome-based management of their portfolios.

Some agencies worked together to develop shared outcomes in their sector, since they see that their outcomes can only be achieved if they worked together. Agencies in the justice sector that were involved with the Pathfinder project pioneered such collaborative work to develop outcome definitions and measures in their sector. Two interviewees from different agencies also described the involvement of other agencies in developing shared outcomes. One agency, which was also involved in the Pathfinder project, pursued this approach only in some areas where there was a function

overlap. Another agency recently set all its outcomes as shared outcomes, working closely with other related agencies in setting targets and planning its contributions.

Officials from central agencies, consultants and auditors were cited as other actors involved in the development of the components comprising outcome information. Respondents from five agencies stated that they asked for central agencies' involvement both in the early development and in the later revision phase of outcome framework and measures. They argued that the tools and guidance that were available for them were not clear enough, so it was necessary to include officials from central agencies, whom they assumed were experts in the area of outcome management. Two agencies hired consultants to help them during the process, especially in understanding the intervention logic aspect of the outcome framework. Auditors joined in the development in later periods, when the OAG initiated a review of the quality of performance reporting (Controller and Auditor General, 2008). Thus, typically the involvement of auditors and central agencies was described in an interview as follows:

The main measures of the department's performance are based on the outcomes framework. These measures are developed using a departmental-wide approach and in collaboration with the Office of Auditor General, the Treasury and the State Services Commission. (Respondent from Agency C)

This study's findings are consistent with the results of the review done by the Economic and Strategy Group (2003). Even though the review did not disclose the various actors involved in the process of outcome reporting, it identified that much of the work of establishing outcome frameworks and specifying outcome measures was done internally within agencies. It was argued in the review that external dialogue involving the presentation of the outcome framework was more for the purpose of communicating it, rather than for gathering inputs to improve the framework.

## **6. Information System for Performance Information**

The level of investment in a new information system that was specifically developed to provide performance information for accountability documents was very low. This was not a surprising finding, since the level of external pressure to reveal information was found to be very weak. The incentive was not strong enough to justify such investment. Agencies tended to use data that was available internally and adjusted the information system over time to provide more appropriate measures.

Since auditors generally checked an information system to determine whether it was robust enough to support the output measures reported in the audited Statement of



Service Performance, agencies only reported the measures that were available in their system. One of the processes that an agency used to ensure that data was available for audit purposes was requiring everybody in the agency to confirm the availability of data in their system. An interviewee described this process by saying:

We require everybody within the organisation ... to confirm at the start of the year that they have information that will support that [output] measure and we do it [through an] internal audit process to make sure that we just get the second report back to the Minister, ... and that's two three weeks. We have said to these guys, 'you said that, prove it to me, show me the documentation.' (Respondent from Agency D)

Respondents from two agencies found that their data was not consistent over time and/or across the organisation, and they tweaked the information system to enable it to produce consistent performance measure data. The database tended to be maintained locally, except in the case of agencies that had offices in various locations around the country and/or around the world; in such cases the databases were maintained centrally. In the case of policy advice output production in which the database was not available, an agency relied on internal periodic reporting as the source of data for its accountability documents. There was one case of policy advice delivery where an interviewee stated that they maintained the database only for accountability reporting purposes, and not for any internal use.

For outcome information, the approach for data availability was different. Since the information was only reviewed by auditors and not audited, the pressure to invest in the information system for outcome reporting was much weaker. Only two out of nine agencies interviewed were able to provide some measures on outcome information. One agency had already adopted its outcome-focused management system some years before the MfO initiative was rolled out, so it already had an information system that was able to produce such outcome information. The process of developing new measures was directed more to internal needs, as one interviewee put it:

...from time to time we develop new indicators where we think that there is a value of doing that and a good reason for it... So, yes and similarly with performance measures, often you know you have a trap if you put in measures just because you [have] got the data to do them... these are two cases where we have developed from scratch. (Respondent from Agency E)

Thus, only if the agency's information system was able to produce outcome measures for internal uses was this information reported for external purposes.

All agencies had their own units to process data for performance reporting purposes, as well as the units that usually processed financial information. These separate units collected the data and analysed them for the purpose of preparing accountability documents. Some bigger agencies used such units to provide internal performance assessments, especially for units or employees responsible for service delivery outputs. Smaller agencies combined the production of financial information and non-financial information in one unit and hired consultants in some cases to help with difficult sections, such as outcome framework development.

## **7. Motives for Providing Performance Information**

At a basic level, agencies were reporting performance information because of the accountability requirement. To the extent that they were able, they would provide all information required to be reported in the accountability documents. Unfortunately, the regulations specified the information requirement for outputs and outcomes in general terms only, so there was a large degree of discretion available to agencies when reporting the required information. All organisations disclosed the information required in their accountability documents only if they already had it in their information systems developed for other purposes, such as budgeting purposes or internal uses. This was very different from the reporting behaviour seen when agencies developed an information system solely for the reporting of financial information for external purposes.

Getting a clean audit opinion provided another pressure, but for unaudited information the pressure forcing agencies to improve the quality of information was very low. Even some mandatory information, such as cost-effectiveness measures, was not reported properly by eight out of nine agencies targeted for interviews, since it was not part of the information required for auditing. One interviewee whose agency recently provided such measures raised his concern on this issue:

There [are] only two statutory components: your financial and your service performance information. Everything else is optional. Our outcome measures in the front of our two annual reports, I [have put] them in there because we have done this work and because we need to [include it], because [of the] statutory accountability to report on the end outcomes. [But], as far as I can see nobody has been trying to report it. (Respondent from Agency G)

Improvements, if they ever occurred, were slow and piecemeal, following the yearly audit cycle. If improvement was made in one period, it could only be incorporated into the next period. One interviewee commented:

So, [there are these] fixed points in time when you can make improvement[s] so if you use your SOI for a particular measure, 18 months later you had to report against it even if you have made significant improvement[s] in the measures you want to use along the way and in the intervening time. (Respondent from Agency I)

Agencies would disclose more, if they had other motivations to do so. One such motivation would be to enable them to get more funding. When a situation is conducive to making requests for more money, an organisation is motivated to produce more performance information to support its request. For example, when a new minister assigned to a department is a strong political figure in government, his/her department uses this occasion to obtain significant increases in funding, as has been discussed previously in the section regarding ministers' interest in performance reporting. These ministers would be willing to invest more in the system to produce such information. In discussing the funding motive, a top-level official from one such agency revealed that:

The ability for the system to engage [in developing] a framework of outcomes ... is reasonably helpful and it does put people whose professional upbringing and mindset is not in that public service ... into a shared debate here. ... Of course, most of the debate is about extra [funding] ... and better access to service ... one of the most difficult bits is having a debate about what are we going do less. (Respondent from Agency A)

In other cases, an independent underperforming agency was included in another ministry because of reorganisation. The ministry was able to secure more funding for this ailing agency and its accountability documents were used as a part of the negotiations with the select committee for that purpose.

The presence of the funding motive ensured that some agencies' activities already had their outcomes set in the budget process. Agencies were likely to report performance information if outcomes were established during their budget process. Even though there was no requirement for ministers to provide outcome measures in support of any new proposals, there was a requirement to indicate how the intervention indicated in the proposal would be evaluated. Treasury might negotiate the criteria used for evaluation, so there was an external pressure on agencies to provide appropriate criteria/measures for such evaluation. Thus, in such a case, as Ussher & Kibblewhite (2005) argued, outcome goals drove the budgeting process which in turn drove the

management process of intervention delivery. Unfortunately, this evaluation process was not that clear for the existing (baseline) expenditure, resulting in less incentive for agencies to develop outcome measures.

For small agencies whose existence was at political risk, there was a strong motive to use accountability documents for public relations purposes. This motive was also strong for agencies where the public perception of them was unfavourable, as a top-level official in a small agency stated:

One of my criticisms with the MfO thing is that in the end [it encourages] you ... to tell [a] beautiful story that actually does [happen, rather] than link [it] to your influence in the work that you do.  
(Respondent from Agency F)

Reputation was also a strong motive for putting more effort into and paying attention to the quality of reported information being provided. There was a case in which a minister used the publication of performance information as an incentive to improve the work of his organisation. He periodically asked some regional units in the organisation to publish some performance measures, where the units were able to provide information to back up the performance. These requests certainly gave added incentive to regional managers' efforts to achieve results and focused attention on the performance information produced. As one interviewee put it:

...the incentive mechanism he [uses] here is shame. It is a shame based model. So, he has thought (about) the problem, [he] cannot manage [the] chief executive pay ... [so] ... he has used shame plus competition, [where] the ... organisation [is made] to feel sensitive that they're at the bottom of some arbitrary rank. (Respondent from Agency A)

Respondents from two agencies also discussed the implications of poor credibility and reputation when agencies received a qualified audit opinion or an unsatisfactory review from auditors regarding the information in their accountability documents. It was necessary for agencies to report at least to the level that was satisfactory to the auditors, so as not to receive negative attention from the minister and members of the select committee.

Concern about reputation is strongest where certain individuals, who are either appointed as chief executives or who are employed as those responsible for preparing accountability documents, are well known in the New Zealand public sector environment as experts in the area of outcome-orientated management system. Mr. Mike Byers is one such individual. He is known for his involvement in many organisational and governmental reforms, including New Zealand public sector reforms

in the late 80s and early 90s. He was then appointed as the first chief executive of the newly established Department of Corrections, and was known as one of a few chief executives who were pioneering the application of an outcome-oriented management system in the New Zealand public sector. Because of his leadership, the Department of Corrections was praised for its initiative in its reporting practice for output and outcome information, even before the MfO initiative was rolled out. Some of his team members at Corrections were later either involved with the roll-out of the MfO initiative or employed in some agencies pursuing better performance information reporting. In an interview, respondents from one agency stated that they were referred to these people when they asked for help in dealing with difficulties in preparing their agency's accountability documents.

These individuals were willing to take risks and put more effort into pursuing an improved reporting practice in their respective agencies. One of two respondents who worked with Mr. Mike Byers for a while recalled that he was given a secure environment for experimenting with ways of improving performance information in managing the organisation. Their organisation invested a lot of time and effort in pursuing better performance reporting. In one case, for example, they were given three years to find ways to produce a performance reporting system that enabled them to link their resources and the targeted outcomes chosen. Another respondent, who is now the official responsible for developing accountability documents in another agency, pressed the top management for more resources to be put into his efforts to develop cost-effectiveness measures. He claimed that this investment had made his organisation one of the few that was able to report on quantitative measures for cost-effectiveness in its accountability documents.

## **8. Disincentives for Revealing or Improving Performance Information**

The issue of controllability has led to a sense of uneasiness in agencies when reporting outcome information. There was a concern that controllability drove risk-averse behaviour that later killed off many initiatives that potentially would improve outcomes. Putting outcome information into accountability documents produced more of an 'under-promise and over-deliver' type of behaviour in public sector officials, even though they were not responsible for the outcome. The main factor behind this behaviour was the tendency of the competition, in the political setting, to asymmetrically value bad news much more than good news, as bad news was more likely to be used in order to score a political point in politicians' parliamentary or public

debate. An accumulation of bad news made it easier to change the public's political perceptions than an accumulation of good news. As outcome achievement was not entirely under the control of an agency, the likelihood for bad news to gain prominence was higher than that for achievements that were more controllable, such as the case of achievements from output delivery. Thus, it was best for officials to under-promise. The best way to under-promise was, as far as possible, never to promise anything, or in the case of performance reporting, less reporting was deemed to be better because then there would be less need to promise to deliver. One top level official explained the under-promise tendency as follows:

The other thing is [with] a lot of public sector on the ground up measures, they measure the failures, they measure road deaths ... And that is part of the function of media ... Ministers do not want to hang their heads [in shame because of] achieving something [negative that] may not happen, and public servants are also very nervous about promising things that they cannot necessarily deliver, particularly with the linkages that get a little... tenuous ... I can probably deliver ten percent improvement but I will promise five because that way I can be sure of meeting it ... by the way that is going to give me something [in the back pocket] ... for the following year .... (Respondent from Agency A)

This asymmetrical value of information also discouraged agencies from experimenting with alternative activities to find better ways to achieve certain objectives or outcomes. Riskier strategies were pulled back from as soon as some short-term indicators showed unfavourable results, not because the intervention failed, but because the data was not calibrated properly to measure the intervention. One interviewee described his experience in this situation:

But because of this [unfavourable] event, you do get a lot of negative attention .... And, that then has a deleterious effect on you as an organisation because everyone starts pulling things back and so the original strategy like getting people out [to] work or whatever was compromised because of the risks. It is sometimes not because [the risk has] increased but [because] the risk could be increased. (Respondent from Agency C)

The tendency to limit performance information also occurred in politically sensitive areas. Some agencies did not even disclose their biggest achievements because they were considered to be very sensitive in the New Zealand setting. For example, some public agencies competed internationally with similar public agencies in other countries. They did not want to disclose some performance measures in case that would in some way reveal the organisation's strategy. In another situation that is politically

sensitive, one respondent stated that his agency did not disclose one of their biggest achievements because it was politically controversial with the New Zealand public.

## **9. Major Obstacles, Constraints and Difficulties**

Some obstacles or constraints were raised by respondents, regarding the process of providing better performance information. There were two types of performance information that in their opinion were difficult to report; the first was outcome information and the second was measures for policy advice output. Most of the difficulties in reporting outcomes were centred on the complexity of the information. On the other hand, difficulties in reporting policy advice related to the selection of proper measures that could be consistently applied to a variety of policy advice outputs which may vary from a simple one-paragraph opinion to a thick analytical report.

In developing performance information, respondents from six agencies argued that outcomes were too complex to be reported. The relationship between outputs and outcomes was more complicated than the relationship of inputs and outputs, where the relationship was seen as linear. One interviewee said:

The linking [of] those to outcomes ... proved to be difficult ... we went through a phase in the public service where people ... [were] drawing nice little middle alliances between output classes and appropriations. ..., I just think that is meaningless. I think that it is a problem but it is also strength in that there is not [a] one to one relationship. (Respondent from Agency A)

This complexity was also related to the respondent's view that so many factors outside their control contributed to outcome achievement. There was a story for every eventuality. The challenge was how to understand these factors, measure them and incorporate them into the report. But, since agencies were not responsible for outcomes and the principal's pressure to deliver this information was low, there was a lack of initiative to overcome the challenge. One top-level official commented:

We did appreciate they [outcomes] were important tools and an important way of communicating to the public and to ministers and other stakeholders of what we [were] about, what we have, so we have always made our best [efforts] to try and make [the information] as meaningful and useful as we [could] but it has been very difficult. And, perhaps ... maybe we could've done more. (Respondent from Agency B)

This situation led to the tendency not to allocate enough resources to performance monitoring systems, as acknowledged by one of the executives of an agency interviewed, who stated:

Yes, the challenge for us as a ministry, [is] we have not had any resource allocated [to] monitoring performance. It was just [a] deliberate choice of the chief executive and minister of the day... We spent a large amount of our vote doing [service delivery], but system monitoring we do not do and outcome monitoring we do not do. (Respondent from Agency D)

Another difficulty related to the problem of measurement. This problem was more apparent when measuring the outcomes of policy advice output, with a dilemma evident when deciding between using short-term measures or long-term measures. When measures that cover a long-term horizon were selected, they did not change much, so it wasn't possible to tell whether something had actually happened and in which area improvement had happened or whether the outcomes had contributed to any improvement. In addition, these long-term measures were very costly to produce if they were not available as part of national statistics data. The measures chosen should show the impact of any changes in society or in the environment. One example was seen in the case of the Ministry of Fisheries. To produce annual data on fish stocks was not feasible under the current level of the ministry's funding. If short-term outcome measures were selected, these could better explain any achievements and the agency's contribution, but the measures tended to fluctuate, as the outputs varied from period to period. Consequently, consistency of reporting was quite difficult to adhere to in this environment. As one interviewee put it:

So, [Agency G] is a little bit [more] dynamic than a lot of agencies because we sit at the top ..., we are [an] economic development agency, we are the environmental agency, we are a people-centred [helping] agency as well, ... ..we [have] got all [these] interfaces, so the environment is quite dynamic, ...therefore your output set is changing as much as your outcome set is changing. So that is one of the things that I have been struggling with. (Respondent from Agency G)

One respondent further explained that they faced a dilemma in respect of their role in selecting the appropriate horizon for the outcome measurement of policy advice outputs. On the one hand, ethically they had to serve the public by giving independent, honest and fair advice to the minister. In their business, there was a trade-off between selecting outputs that would provide immediate benefits but would not improve long-term outcomes, or of selecting outputs that might not provide benefits in the short run but which had a greater probability of producing better long-term results. Therefore, if they acted ethically, they would select a longer time horizon for outcomes and also select one set of outputs. On the other hand, they also needed to meet the minister's requirements by providing the kind of advice that he/she favoured. In this case, the



selection of the outcome horizon tended to be shorter, following the election cycle of a three-year term. This in turn resulted in selecting a different set of outputs. Furthermore, in some cases in which outcomes did not change much during an election cycle period, the minister preferred a strategy (set of outputs) that was visible to the public in the short run even it might not be as effective in achieving better long-term outcomes. For example, one interviewee explained that based on his work experience, increasing the number of enforcement officers was the least effective intervention in the long run to reduce regulation violations. However, it was always the preferred intervention for the minister, as it would be easier for him to show that he had done something during his administration. Thus, agency officials were sometimes forced to provide interventions that they knew would not work to change outcomes in the long run. Furthermore, they were criticised by auditors for selecting inconsistent measures of outcome in order to comply with their ministers' interests.

Measurement difficulty also arose in cases where the performance of policy advice delivery was measured. There was a sense of ambiguity from some interviewees on the appropriateness of reporting policy advice as an output. Respondents from three agencies argued that policy advice was just an intervention or an activity that was a part of what was actually a real output. In these cases, policy advice was an input into the bigger activity of policy delivery. One interviewee stated that there should be a distinction between a set of extensive actions that resembled a policy development (Big P) and a simple activity that provided mere comments to the minister (Small P). He further argued that the Big P was the real output, while the Small P was not. However, another interviewee argued that in some cases the impact (outcome) of a one-page comment letter was bigger than a one-hundred page report of an extensive piece of policy research, so the obvious size difference between Big P and Small P could be misleading, in light of their differing impacts. This debate highlighted one important characteristic of policy advice that showed how it was difficult to develop performance measures for output delivery. One interviewee commented on this difficulty:

If you write a ministerial [report] in response to a particular letter that might raise four or five different things ... so is [this] a single output or is [it] a response... to a particular project that might [be] referenced to that ministerial [report] ... These are [some of the] debates about what [is] the best unit of measurement. (Respondent from Agency A)

Policy advice outputs varied widely from one output to another in terms of inputs, activities and impacts. Reporting them in standardised measures was meaningless, since the measures were not those that drove the agency's business to deliver policy advice

outputs. One respondent stated that what drove them to manage policy advice delivery was how well the policy advice influenced decision-makers or people, but measuring this influence was not an easy task.

The early review of the MfO initiative by the Economics and Strategy Group (2003) also showed the difficulty of specifying performance measures for policy advice agencies. There was an ambiguity evident in choosing between long-term and short-term measures. The review quoted a Treasury official who proposed measuring the impact of policies as outcome measures for policy advice, but concern was raised about the complexity of the link between intervention and impact.

Another challenge was being able to come up with cost information at an individual output level. Currently, cost information was available at output class level, since it was developed as a part of the budget process. One interviewee, who did data analysis for unit performance reporting, agreed that cost information was not adequate to allow his agency's ministers or stakeholders to assess efficiency at output level. Management demand for a costing system was not strong enough to justify a substantial investment in the system. This challenge was described by this interviewee as follows:

But at output level, one of the [difficulties is that], we have our outputs trying to match up [with] our financial [information. To] be able to cost our outputs is quite ... challenging in our organisation. And the reason for that is that ... our staff on the front line ... do not have a time recording system. ..., [This] is one [issue] that I am getting [on to], and we [will] do it based on time estimates. (Respondent from Agency B)

Agencies also had to manage some degree of disagreement in the process of developing an outcome framework and measures. The process involved people across an organisation, each of whom would have different interests and priorities, and therefore it was common to have some disagreements along the way. The disagreements seemed lower in agencies that focussed on policy advice delivery. These agencies used a top-bottom approach in developing outcome frameworks and measures, so everybody was forced to follow what was demanded from top management. There was strong leadership from top management for the development of an outcome framework for policy delivery. On the other hand, for agencies with a service delivery focus, the disagreement was stronger, since they used more of a bottom-up approach during the development process. An interviewee representing one such agency commented:

It is ... very difficult because we are talking to different people across the organisation that do totally different things and they all think that what they do is the most important thing.. (Respondent from Agency H)

Strong leadership in top management positions was cited as the way to resolve this disagreement. Interestingly, one interview revealed that a service delivery agency that used a combination of top-bottom and bottom-up approaches during the process had a low level of disagreement when developing an outcome framework.

Another impediment was a lack of top management's interest in outcome-focused management. In these agencies, the progress in adopting outcome reporting was slow. Respondents from four agencies that had this kind of top management for some time periods during the MfO initiative argued that the auditors' pressure via a management letter had led to a gradual increase in resources to implement a rolling review of their management system to improve the quality of outcome information. When an agency had a chief executive that had already bought into outcome-focused management, progress was faster. As has been discussed before in relation to the role of reputation, the early adopters were the ones that had been known in the New Zealand public sector environment as high reputation career bureaucrats. Two respondents who had previously worked with one of them argued that his chief executive officer already had an outcome focus in mind, and therefore was willing to put more resources into his agency to produce outcome information, even before the MfO initiative was rolled out. His agency was among the top scorers on outcome information reported during the period of observation studied in this thesis.

Some of the factors mentioned by interviewees that contributed to the slow progress in outcome reporting were related to the role of central agencies and auditors in the development of performance information. The first factor was the view that the system was being forced into an implementation phase without considering appropriate incentives and support. As one interviewee put it:

They put the requirements [from] PFA to report and ... become audited [into place] but they have been ... slow ... [in] providing the guidance, forcing every public agency to invent the wheel... I think [it was] an outrageous imposition of ... cost ultimately when as central agencies, they should work [on] far more robust methodology, rather than just do it by ramming it through and making it auditable. (Respondent from Agency D)

In terms of support, the main thing that was provided was reporting guidance, but it was not adequate to help agencies produce reports to a level that was satisfactory to auditors. All respondents stated that the available guidance did not help when developing outcome information. The guidance was very simple, so it only fitted the purpose of getting a basic idea of the information structure required. The respondents wanted

specific examples that were applicable to their own agencies to enable them to develop a specific framework and measures. One interviewee described this situation as follows:

We constantly asked them [for] the best practice examples, but they do not actually have it ... what I like is [for] them [to]... provide a case study on 'this is how you do it.' Because we do involve them early and we involve Treasury and audit and lots of our staff early [on] but we [found] that .... that they do not [have] that information [available]. (Respondent from Agency F)

The second factor was the agencies' frustration regarding disagreement about the acceptable level of performance information. Respondents from six agencies described their frustration with the different views held by central agencies and auditors on how to report some information, since the reporting guidance was very general and broad. The focus of the disagreement was in the area of outcome and cost-effectiveness measures. One interviewee described this situation:

The other thing was of course we have had problems over the last couple [of] years because Treasury and Audit have different views of what should go on the SOI. Treasury have one view and Audit have another, which ..., from a departmental point of view, [made it] very difficult to work out what you put in... Nothing in the guidance from ... Treasury would lead [you] to produce [what was needed] ... we actually still do some much more effective [ways] of telling [the] story of what [we] are wanting to achieve. It is pretty high level and the audit stuff is pretty low level. Somewhere in the middle is actually what we have to produce. (Respondent from Agency F)

The third factor, which stemmed from the absence of any specific reporting guidance, was the lack of useful feedback for agencies to be able to improve the quality of their performance information. Periodically, agencies received comments and criticism on the reported information, but the comments were general, and in the words of one respondent, were "very theoretical". A comment from an interview described this situation as follows:

During the audit process, it becomes a matter of judgment. They say we are not satisfied with this [but] they generally do not tell you why they are not satisfied with [it] and how you can improve it. What we want from them is a way in which we can improve. (Respondent from Agency D)

Auditors argued that their statutory responsibility was to assess the information. It was the responsibility of the individual agency to find out how to report it properly. Unfortunately, agencies did not know how to report properly and the central agencies that provided soft regulation on this issue were not helping them either. The reporting

system for non-financial performance was not as sophisticated as the system for financial information where a detailed standard were available for reporting, so obviously a greater level of disagreement existed among preparers, regulators and auditors on how to report non-financial information. This situation created a high degree of frustration for preparers, as they were frequently criticised but were unable to address the criticisms properly.

## **10. Performance-Based Compensation**

Respondents generally expressed a favourable attitude toward output-based compensation, but the level of support was not high for a compensation system that was based on outcome performance. Three agencies based some of their employee compensation on the achievement of output and short-term outcome measures. One agency tried this compensation scheme for its employees, but the minister did not like it and so it was abandoned. Almost all respondents argued that if performance-based compensation was to be applied, it should include some consideration of the effectiveness of the interventions. One respondent stated that quality measures could be a proxy for effectiveness, since they exhibit some aspects of effectiveness, especially in service delivery outputs. Others supported the use of short- or medium-term outcome measures in policy advice areas, but with some degree of judgment being needed.

Respondents from only two policy advice departments had misgivings about performance-based compensation, because of controllability concerns. One of them stated that:

I would only tie performance to something which I thought the person had control over. [If] they do not have control over it then that [outcome] is nonsense. (Respondent from Agency F)

He further argued that performance-based compensation would further encourage agency behaviour of producing policy that the minister liked, rather than giving him or her fair and frank advice.

## **C. Discussion**

Discussion of the findings from disclosure analysis and the interviews is presented in five interrelated topics. The discussion starts with two sections of general analysis, using the theoretical framework selected and developed. This analysis provides an overview of the environment in which reporting practice takes place. The first section of the discussion deals with the analysis of public choice problems in the

New Zealand public sector system, by critically examining the Westminster system. The second section presents the analysis of the principal-agent relationship in the New Zealand central government setting. The third section focuses on an evaluation of Managing for Outcomes, using the perspective of public choice theory and agency theory. The fourth section discusses the issue of incentives for agencies to improve the quality of their performance information in the New Zealand setting. The last section deals with challenges in the reporting of performance information.

## **1. Public Choice Problems in the New Zealand Public Sector: An Analysis of the Political Constitution of the Westminster System**

The political system in New Zealand is modelled on a form of parliamentary democracy developed in the United Kingdom, which is known as the Westminster system. Rhodes (2005) described four characteristics of this system of politics. First, parliament, with its unity of the executive and the legislature, is sovereign. This characteristic is also called the doctrine of parliamentary sovereignty. Second, political power is concentrated in a collective and responsible cabinet as the executive component. The third characteristic is that ministers are accountable to the parliament. This characteristic is the complement of the second characteristic. Together, they are known as the doctrine of ministerial responsibility. The last characteristic is that a constitutional bureaucracy should be non-partisan and employ expert civil servants. This last characteristic is also known as the doctrine of political neutrality of bureaucracy.

Cowen et al. (1992) added some more characteristics to the above description of a typical Westminster system of politics. They note that there is no written constitution in this system. What acts as a constitution is a combination of statutes, procedural rules and conventions. They argue that party discipline is fundamental in the conduct of parliament. Political parties will use their “constitution” and “punishment” to facilitate party unity in parliamentary voting on legislation. The last characteristic is the presence of a plurality rule or winner-takes-all rule, in which a member of parliament is elected from single-member districts, based on the highest number of votes in the district. In the New Zealand political setting, this characteristic was abandoned with the adoption of the mixed member proportional voting system (MMP system).

It is important to consider some of these characteristics when analysing the public choice problems found in the New Zealand public sector, because they serve as the rules of the game for the behaviour of political actors and bureaucrats. In addition,

the change in the voting system to MMP should also be integrated into the analysis, to show whether the problems are mitigated or magnified by the new system. The extent to which public choice problems are found in the political market is likely to be shaped by a combination of the doctrine of parliamentary sovereignty, the doctrine of ministerial responsibility, party discipline characteristics and the election voting system. On the other hand, the extent to which public choice problems are found in the bureaucratic hierarchy is likely to be influenced by the doctrine of ministerial responsibility and bureaucratic neutrality.

### **Public Choice Problems in the New Zealand Political Market**

Somers (2008) argued that the doctrine of parliamentary sovereignty can be defined through two perspectives: the legal perspective and the political perspective. In the legal perspective, the sovereignty of parliament means that only parliament has the right to make or unmake any law and no other person or body is able to override any law. In the political perspective, sovereignty means that only parliament is the source of authority for government, since it is the only institution of the central state whose members are chosen by the people. The majority party or coalition in the parliament forms a government by constructing a cabinet which consists of its elected members. Cabinet members, who have the authority to control bureaucratic organisations, are collectively and individually responsible to parliament. Therefore, parliament represents the people who have the authority to subject executive and bureaucratic organisations to be accountable to them. In other words, parliament has the authority to scrutinise government policies and operations, using either its legislative and financial control mechanisms or through other available opportunities, such as select committee systems or employment of parliamentary officers to help them perform such scrutiny<sup>3</sup>.

In public choice theory, this sovereignty is the source of rents in the political market. Any parliamentary legislation or parliamentary-controlled bureaucratic actions may involve involuntary transfer of wealth from one group of people to another group in society. Political decision-making is the means of allocating rents in the political market. The types of public choice problems that may occur depend on the likelihood that politicians are influenced by interest groups or their local constituents and the nature of the mechanism used by the party/coalition to ensure that its members vote support for the party's/coalition's preferred policies (party discipline).

---

<sup>3</sup>For example, parliament employs the auditor general to assist parliament in scrutinising financial matters of government and employs ombudsmen to assist parliament in the investigation of complaints against government departments.

The change in the voting system to MMP to elect parliamentary members in New Zealand has changed the pattern and the degree of special interests' influence. In their analysis of the proposal to change voting from the first-past-the-post (FPP) method to the MMP method in New Zealand, Cowen et al (1992) argued that the change would affect the influence of interest groups and their local constituents. The proportional representation method weakens the influence of the local constituency on politicians, since some of them can be elected even though they do not win in their district. Furthermore, they rely to a greater extent on their parties to gain seats during an election. Consequently, their parties are able to exercise stronger party discipline. However, the effect of the voting change on the influence of interest groups is not clear. Interest groups have to channel their pressure through multi-party coalitions, instead of through the majority party as was the case under the previous voting method. With MMP, the tyranny of the majority, in which the majority imposes their preference on the minority, may be weakened. In a coalition, the majority party has to give concessions to minority members to secure their confidence vote.

There should be mechanisms by which the discretionary power of a politician is limited with regard to increasing the amount of rents, minimising log-rolling and pork-barrel projects in the budget allocation. Politicians need only consider voters when making their decisions if there is no constitutional mechanism to control their legislation or actions. Uninformed voters may not be able to provide adequate disciplining of the behaviour of politicians. This means that parliament itself controls the behaviour of politicians in relation to political decision-making. This self-control and the re-election motive may not be strong enough measures to mitigate political capture when opportunities exist.

The current discretion limit in political decision-making may not be enough to mitigate public choice problems in the New Zealand political market. Currently, there is only one discretion limit required to be met in the financial decision-making carried out by parliament; that is, there is a requirement for cabinet to apply principles of responsible fiscal management in pursuing its policies. The government's financial information requirements help parliament and the public to assess a politician's behaviour in the cabinet as to whether this financial discretion limitation is followed. This limit may control the total amount of rents that can be allocated by stopping the trend of excessive budget increases, but it may not be able to reduce the rents that are available to be allocated from the existing/current budget. There should be a similar parliamentary and public scrutiny for this budget. One way to do this is by requiring



politicians in the cabinet to state their objectives (outcomes) in the budget allocation and later requiring that outcome information be produced, so that progress towards achievement of outcomes can be assessed. The MfO provides the latter component, but unfortunately there is no effort to address the lack of the first component.

The necessary complement of parliamentary sovereignty is the doctrine of ministerial responsibility. Ministers in the cabinet act as a bridge between parliament and the Crown on one side and its bureaucratic organisations on the other side. Therefore, they are collectively and individually responsible for the conduct of the government through these organisations. Ministers are responsible for securing and providing information to the cabinet and parliament, so that parliamentary scrutiny can be implemented. In addition, ministers should also take responsibility for the actions of the bureaucracy, since these actions involve the political/power dimension, which means they are subject to ministerial control for the political legitimacy of these actions (Somers, 2008). From this doctrinal perspective, since the parliament's power to control the bureaucracy has been delegated to ministers, they must bear ultimate responsibility for the actions of the bureaucracy, even though they may not have direct control over them.

The ministerial responsibility doctrine is implemented in a number of ways, but these are less clear than what is laid down by legal requirements. At a minimum, ministers should be answerable for personal and departmental actions. This is the first thing they will do, because it does not cost them much compared with the possibility of allowing an adverse public perception of them to develop. For bigger problems, addressing them in front of parliament may not be enough. Ministers may be forced by their colleagues or the public to show that they are taking all necessary steps to rectify any mistakes. Lastly, ministers may be pressured to resign by their colleagues in parliament and/or face dismissal by the public at the ballot box. This last step also depends on their colleagues' pressure on them to resign or to stay and the public satisfaction with the minister's explanation or corrective actions.

The most important feature of this doctrine is that its operation is all political. When ministers are designated to be responsible for outcomes in a reformed public management system in New Zealand, this responsibility can only be carried out politically. Attempts to put such responsibility into a statutory requirement, for example by requiring ministers to report outcome information, will not happen. Clarifying ministerial responsibility may provide more certainty with the assessment of outcomes, but as Somers (2008) argued, such clarification may restrict the scope of responsibility

held because there should be a causal notion of responsibility. It may also be difficult to hold ministers responsible for lesser offences, because the bar of assessment is higher now as a result of clarification. In the case of the New Zealand public sector, the PFA states that ministers are responsible for outcomes and chief executives of public agencies are responsible for outputs. Can parliament and/or the public forgo the claim for ministers to be responsible if their departments cannot deliver outputs that have been promised? It is highly unlikely. Ultimately, ministers will be responsible for any bureaucrat's action, because this responsibility is political.

However, there should be a mechanism by which the bureaucracy is directed to pursue certain outcomes and ensure that outcome information is available, so that the achievement of progress toward outcomes can be assessed. Even though the minister's responsibility will always be political, ultimately ministers are responsible for the outcomes of the bureaucracy. Parliament and voters in the political market need information in order to assess this ministerial responsibility. The MfO initiative went in the right direction in requiring such information, but the pressure was wrongly directed toward public agencies to provide it as a part of their accountability documents, instead of pressuring ministers to declare that the outcomes reported are their responsibility. The information requirement prescribed under the MfO initiative does not address the real accountability requirement for outcomes.

### **Public Choice Problems in the Bureaucratic Hierarchy**

The interplay between the doctrine of ministerial responsibility and the doctrine of political neutrality influences public choice problems in the bureaucratic hierarchy. The latter doctrine is also part of the established New Zealand political system. The origin of this doctrine in the New Zealand setting can be traced back to the enactment of the Public Service Act 1912. There are several components of this doctrine. First, political interests should not influence employment decisions. Second, public servants should serve the government of the day professionally and impartially. Third, public servants should work equally well with all governments and ministers that come into power, no matter what their political persuasion is (State Services Commission, 2004).

Somers (2008) argued that this doctrine is about political control of parliament over bureaucracy. This convention makes possible parliament's control, via any government of the day, over the bureaucracy in order to pursue their preferred policies in the delivery of public services. Bureaucracy has to be non-partisan because it has a political and moral duty to respect democratically elected institutions of democracy,

including any government in power. However, this doctrine does not imply that bureaucracy should be neutral towards the government and the opposition. In Somers (2008)'s words, bureaucrats should be political chameleons, who surrender political power to elected politicians, instead of political animals, that exercise political power independently. Public officials should be politicised in a way that is consistent with political decisions made in parliament, which are legitimised through the democratic process, rather than by exercising their own private morality, ethics or partisan values.

The imperfect functioning of the political neutrality doctrine is the essence of the public policy problem. By pursuing their own interests when they differ from the interests of the political principals, bureaucrats essentially violate this doctrine. Ministers, in following the doctrine of ministerial responsibility, should be able to control the actions of supposedly politically neutral bureaucrats. Therefore, public sector reform was initiated to put the bureaucrats under the control of their political principals so that they would be politically neutral when serving any politicians in power.

This control was asserted in the reform process by specifying the relationship between ministers and chief executives and ensuring that there was a clear responsibility laid on each of them. The New Zealand model is quite clear about these responsibilities conceptually. Unfortunately, the model seems to ignore the fact that the ultimate responsibility for the conduct of public officials is in the hands of ministers. Thus, the accountability of chief executives cannot be easily separated from the accountability of ministers in managing their portfolio of public agencies. The prescription that government agencies should report performance information to fulfil their own accountability requirements is not what is operating in the New Zealand political realm. Ministers are more concerned that the reported information does not show anything that is unfavourable and, in the extreme, makes them answerable or under pressure to resign. A respondent from central agencies described this interlinking of ministers and chief executives in an interview as follows:

The trouble with the ministers is they are accountable for the actions of their department ultimately anyway, even at output level. .. All of their peers start looking at them for solutions and if they fail to provide solutions, their status within the Cabinet is diminished. Ultimately, portfolios become at risk, or conversely the only way a minister can ascend is to be... seen to be squeaky clean and [a] high performer within the political context. So, ... actually they behave very much like the most senior executives in practice. (Respondent 1)

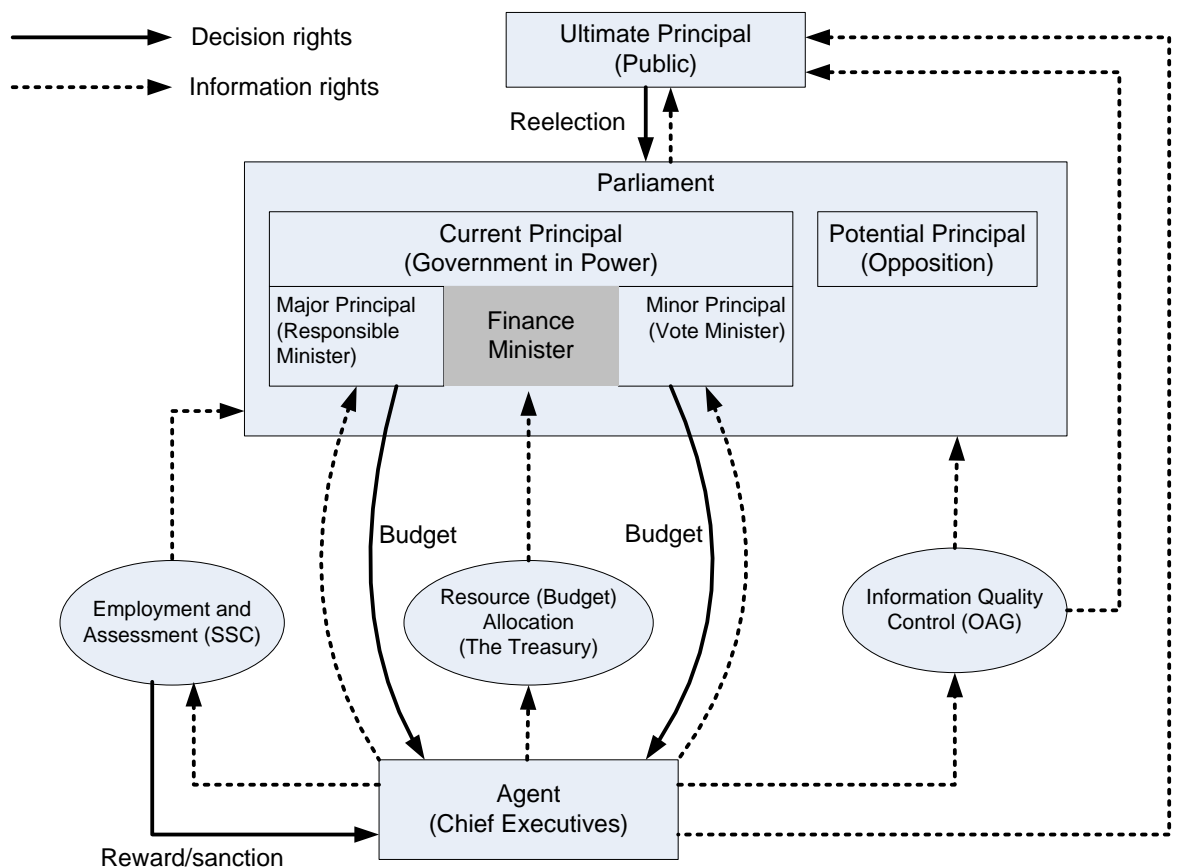
Since the accountability of chief executives cannot be easily separated from the responsibility of ministers, it is difficult to require chief executives to report their performance separately without considering that ministers are also ultimately responsible for this performance. The MfO prescription for reporting outcomes will not produce meaningful information unless there is pressure on ministers (government) to acknowledge that the reported outcomes are also the outcomes for which they are responsible. Ministers should own the outcomes, together with their agencies. There should be a formal mechanism in the budget process for the government (ministers) to state the impacts (outcomes) of government expenditure managed by their portfolio of public agencies and how those impacts are measured. Unless the political principals are pressured to specify their responsibility in the delivery of public services, there is no inducement to provide a proper set of performance information on those services.

## **2. Mapping of the Principal-Agent Relationship in the Central Government Setting**

The principal-agent relationship in the NZ Central Government setting can be simplified as shown in Figure 2 below, based on the accountability arrangement depicted in Figure 1 on page 21. There are four cascading principal-agent relationships. At the bottom of this scheme, the chief executive of an agency is the agent of ministers who have twin interests as owner and purchaser. The chief executive may have two types of political principal: the first one has these twin interests and is usually referred to as the responsible minister, and the other has only a purchase interest and is usually referred to as the vote minister. The responsible minister is the dominant purchaser of the agency's services. In the next hierarchy, ministers can be considered as an agent of the coalition government in power. Then, the coalition government in power is an agent of the ultimate principals, who are the NZ citizens and residents that elect them and give them the task of managing the state.

In this scheme of principal-agent relationships, the opposition coalition can be viewed as the potential agent of the public (potential principal to the chief executives) that is in competition with the elected agent to become the political principal at the next election. The incentive for the current principal (government) and the potential principal (opposition) is the same, how to get (re-)elected. In order for them to compete fairly, public information is provided to both of them, even though the current political principal is highly likely to have an upper hand when it comes to access to unpublished information, due to their power to allocate resources (budget), especially in their interest

relationship as purchaser, and in their leverage in employing the chief executives. Both parties rely on the OAG to verify the public information provided by executive agencies and to advise them on matters related to the information verified. In addition, the current political principal employs three agencies, usually referred to as central agencies, to help them in managing the relationships with the chief executives: Treasury for resource allocation, the State Services Commission for chief executive employment and performance assessment, and the Department of Prime Minister and Cabinet for chief executive coordination (not included in Figure 2 due to their limited involvement in performance information). The potential political principal has information access to these central agencies, but the current principal still retains control of these agencies due to their ability to allocate resources. Since some relationship management activities are performed by these central agencies, they can be viewed in a narrow sense as the pseudo-principals of the executive agencies in relation to these central agencies' functions.



**Figure 2 Principal-agent relationships and accountability arrangements in the New Zealand Public Sector (Adapted from Anderson & Dovey, 2003).**

Information plays a central role in the functioning of the incentive mechanism of the political principals in relation to the ultimate principal. Political principals are interested in the kind of information that might help them to become the elected

principals in the next election. The current political principals will tend to maximise the possibility of good news and minimise the possibility of bad news. The potential political principals have the opposite tendency. Information asymmetry exists between political principals and their agents, so that political principals have to rely on executive agents to provide information. Some information should be provided by executive agencies as mandated by regulations, such as accountability documents. As long as no bad news is reported in these documents, the current political principal will be likely to pay minimal attention to them. Other information is only provided if it is asked for by the current and potential political principals in select committee meetings. The potential political principals may use this access to extract more bad news from the executive agency only if they know what to ask and where to find such bad news. Since agents are controlled by the current political principals, it is safe to assume that they will try to minimise the possibility of bad news occurring and to maximise the possibility of good news. This behaviour will satisfy their most dominant principal. So, even though the executive agencies and the current political principals are at opposite ends of the principal-agent relationship, in terms of information incentives, they are very much aligned.

The findings from the interviews reveal that ministers have become involved in agencies' performance reporting if there was an indication that bad news would be reported. In that situation, ministers' interests aligned with chief executives' interests so that the prescription for information disclosure based on the principal-agent relationship did not work. Internally, in the minister-chief executive relationship, ministers are highly likely to act in their capacity as principals to straighten things out, but externally they have the same interest in mitigating against bad news. Only if bad news cannot be mitigated, such as in the case of public scandal, do ministers externally hold chief executives responsible. Ultimately, internal pressure from their peers in the cabinet and the external pressure of the opposition and the public will make ministers responsible for a bad performance by their chief executives. Therefore, both ministers and chief executives are agents of their prime minister internally in the cabinet, or agents of the public externally, in terms of outcome-focused management and reporting.

The executive agency faces tension between providing information for managerial accountability to the minister and for accountability purposes to parliament. Information for managerial accountability is different from that for parliamentary accountability. The former type of information is produced to enable ministers to manage their portfolios properly. Information is a part of a feedback loop that enables

ministers to learn from past experience in order to manage better in future. Taking risks in reporting is important in this environment, because ministers and their chief executives need information that is operationally useful. The reported information should be able to be used to drive a learning process, in which ministers and their chief executives gain knowledge of what works and what does not work, and enables them to incorporate this knowledge in order to perform better in the next cycle of the policy delivery process. Progression toward better interventions requires trial and error. More information can be provided as long as the benefit of that information is more than its cost.

The latter type of information is intended to fulfil agencies' accountability to the parliament in performing the duties delegated to them, including delivering public services and managing the organisation's resources. Performance information is provided and reported because it is a statutory requirement to measure results achieved against publicised intentions. The purpose is to enable parliamentary oversight. In this setting, the use of information is used to judge agencies' performance. Potentially, there will be some disincentives to providing more information than is required by statute, such as an aversion to incurring media and political reaction to bad news, or uncertainty about the use of extra information by other parties. Providing less information is the preferred choice in this setting.

Ministers need time to understand how to manage their relationships with agencies. There is a bounded rationality problem evident in relation to principals. It is not easy for them to get an idea of what is or should be the outcomes produced by the public agency for which they are responsible. If they know the outcomes to be pursued, they still have to learn what the relationship is between the agency's efforts and other factors that might affect the outcomes. The fact that there is not much involvement of ministers in the development of outcome frameworks, and no meaningful discussion in the Select Committee during budget discussions or hearings, may indicate this situation of bounded rationality.

Furthermore, ministers may have to rely on agencies to a large degree to provide information to them while they are learning how to relate to their agencies. In most cases, agents in the public sector possess an information advantage, similar to the case of the principal-agent relationship in which the agent is an expert or professional. In this type of relationship, knowledgeable agents serve less-informed principals. The information asymmetry can be particularly severe, since principals do not possess the technical knowledge to evaluate the agents' efforts or the outcomes accomplished.

Ministers may have to rely on their own uninformed judgment (Sharma, 1997) in making decisions related to their relationship with their chief executives, such as agreeing on an output plan or supporting an increase in the agency's funding.

Boston (1992) argued that the problem of asymmetric information may not lead to excessive opportunistic behaviour and bureaucratic capture. The budget process provides additional pressure on the disclosure of performance information. This process ensures that performance information is available in the budget, which is also in a department's corporate plan. A performance agreement is developed based on this corporate plan. There is extensive external scrutiny undertaken during the budget process by central agencies, cabinet committees and select committees. This process shows that the specification of performance information is a collective process, in which the minister is but one of the participants. It is likely that chief executives are asked to perform demanding tasks in delivering outputs.

The above argument is valid for output information, but unfortunately the same rigorous process is not available for outcome information. The score results show that the comprehensiveness of output information is much better than that of outcome information. The inclusion of output information in the budget process facilitates a collective effort that helps improve the quality of this information. Nobody in the early periods of the MfO initiative reviewed the SOIs and the outcome information reported in departmental annual reports. Later, the Audit Office provided a review of outcome information, but the review did not have sufficient influence to pressure chief executives to provide more comprehensive outcome information.

Boston (1992) also argued that ministers have an array of supports so they are not wholly dependent on chief executives for information about agencies' performances. This support comes from their own staff, parliamentary colleagues, central agencies, external experts, the audit office and members of the public (interest groups). Thus, ministers may have enough collective skills to understand aspects of their departments'/ministries' business.

However, there is not much incentive for ministers to employ others to help them in dealing with their agents. If they employ experts, the benefits may not materialise in the short run, but the costs are visible in the short term. It is unlikely that experts are able to learn quickly the complexity of the task. In the short run, they are exposed to public and opposition criticism if they spend a large amount of public funds on such advice. The use of purchase advisors for budget scrutiny, by the Minister of Finance in 2009, was subject to public debate on the costs and benefits of such an



approach. If they employ auditors or officials from central agencies, these officials may not have enough expertise to help ministers with their learning process. These people do not have a level of knowledge and experience that is comparable to that of the public officials who run the day-to-day business of agencies.

The only available option is the ability to apply the revelation principle. Agents should be induced to give a full and truthful disclosure of their private information. However, communication restrictions may inhibit the ability to apply this principle. In many cases, the information needing to be disclosed may be too complex, so that it is not easily communicated to or comprehended by political principals. In addition, the adversarial situation found in the political setting may deter public officials from providing more information. Political scrutiny creates an environment in which a long run of good performance can be destroyed overnight by one instance of bad performance. The strategy resulting in this environment is risk aversion. A large amount of time is spent in managing risk so that ministers will not get a 'surprise'.

Political scrutiny also influences the behaviour of ministers and departments/ministries. Chief executives will only report formally on what they have real confidence in being able to measure and prove to auditors. This information may not be that which is used by internal management to tell how well the organisation is doing. The selection of performance measures to be disclosed is more focused on form, with less emphasis on substance. Effort will be spent more on ex-ante accountability, such as preparing budgets and performance indicators, rather than on monitoring of results. The real substance of accountability of agencies to their ministers is exercised through periodic and informal exchanges between them. The formal system of accountability is safeguarded so that no problems will emerge.

### **3. Managing for Outcomes Evaluation**

What the MfO initiative tried to correct is the missing focus on outcomes. The original model was deemed to be flawed because there was no apparent mechanism to ensure that public services were delivered so as to achieve certain outcomes. The first attempt to revise it by incorporating a strategic management framework failed, due to incompatibility of the framework with output-based planning and the unsuitability of applying private sector strategic management concepts in the public sector. The last attempt, the MfO initiative, forces executive agencies to provide outcome information as evidence that they are managing their operations in order to deliver certain outcomes.

There should be accountability for outcomes in the public sector. The elected government constitutionally determines the objectives of bureaucracy. It can be argued that these objectives represent what politicians have promised to their constituents during their election campaign. They have a mandate to fulfil these promises through the achievement of desired policy outcomes. Therefore, it is appropriate to put the accountability for outcomes into the hands of the elected government (and its ministers), which then asks the bureaucracy to deliver the most suitable mix of outputs to achieve the intended outcomes. It is necessary that executive agencies help ministers, since they exist to serve them (Bale & Dale, 1998; Boston, 1992). The MfO initiative has put this obligation on agencies to help their ministers, in its proper place. The missing piece of the puzzle in the initiative is that there is no attempt to make ministers claim or take ownership of the outcomes developed and pursued by their agencies. This may relate to the fact that the MfO initiative is a bottom-up process, so the leadership is just not there.

### **The Tension between Controlling Bureaucrats versus Management Inefficiency of Ministers' Interventions**

The main reason for public sector reform in New Zealand was to mitigate institutional capture. There was a strong conviction that bureaucrats and state-employed professionals, such as teachers, doctors and social workers, had undue control over the general policy that was supposed to be the prerogative of the elected government (Treasury, 1987). The public sector is not responsive enough to the changing preferences of elected politicians, because the incentives are low-powered in a system in which much of the policy process is under the direction of the bureaucratic process. When incentives are not strong enough, the logical step is to move the process into the environment of high-powered incentives. The public sector should be reformed so that institutional responsibility for the direction of policies can be separated from their implementation and administration. The former should be under the full control of the elected government and its minister.

The move to politicise policy direction was taken cautiously. There is a concern that the intervention of ministers in operational details, in response to political pressure on them, may compromise managerial efficiency. The reformers had an ambivalent attitude towards political control and accountability (Mulgan, 2008). On the one hand, there is a desire for politicians to take over political control of policy. However, on the other hand, too much political control will lead to a capture problem in the political market that may diminish the efficiency of policy delivery. This ambiguity led to the

decision to separate responsibility for outputs and outcomes, which may not easily work in practice. Ministers and chief executives must work together to achieve outcomes. The main objective of a government is not delivering outputs, it is about achieving outcomes. What the reform has created is an artificial separation of responsibilities between ministers and departments/agencies and blurred the roles of both.

In an interview, one respondent from a central agency argued that the MfO is not progressive enough to acknowledge this blurriness of roles. He said:

... I have to say that the roles for principal and agent have always been blurred. It did not change with MfO. In some respects, you can actually agree that the principal is now being given the information [needed] which they can [use to] ... perform their role as a principal better. The only better [result] which the blurring might [have caused] is really in making the most senior managers in the institutions equally responsible for pursuit of outcomes. But, in some respect it is a bizarre [situation.] (Respondent 1).

### **The Confusion between Accountability for Outcomes and Accountability for Managing for Outcomes**

There are three difficulties in framing the accountability requirements related to outcomes in the MfO initiative. First, it is not easy to separate accountability for outcomes from accountability for managing for outcomes. Second, there is a difficulty in reporting performance information that can differentiate between these two accountabilities. Third, outcome-related accountability has its inherent difficulties, such as the determination of an appropriate level of outcomes, the complexity found in intervention logic and measurability of outcomes for assessment purposes.

The Economics and Strategy Group (2003) argued that in the original reform, as formulated in the PFA 1989, there was already a requirement that agencies should be “managing for outcomes”. Chief executives were required under the act to give advice and deliver outputs that best supported the government’s desired outcomes. The proper interpretation of giving advice as indicated in this requirement is to effectively determine which outputs best support outcomes. The MfO initiative merely added the requirement to make the outcomes more explicit by pressuring agencies to be actively involved in the selection of outcomes and in determining how to achieve them.

This additional requirement of the MfO initiative has expanded the boundary of the chief executives’ accountability and brought it closer to the degree of accountability required for outcomes. If chief executives are accountable for managing their departments’/ministries’ operations to achieve certain outcomes, there should be a set of

outcome targets available first. When this is not available, and there is no attempt on the part of the government or their ministers to provide it, chief executives are pressured to develop such targets. Ministers' limited involvement in the process of outcome development may indicate that they have distanced themselves from the outcomes reported by chief executives. The outcomes essentially belong to the ministers. However, all respondents interviewed argued that they own the outcomes, not their ministers. Furthermore, most of the departments/ministers referred to these outcomes as 'Departmental Outcomes'. If they own the outcomes, then logically they are responsible for the outcomes, not their ministers. However, the MfO initiative is not intended to take responsibility for outcomes from ministers. Nevertheless, in practice, there is confusion about how to draw the line between accountability for managing for outcomes and accountability for outcomes.

The reporting requirement increases the confusion. Public agencies are required to report outcome information in their accountability documents. If these documents are intended to fulfil accountability requirements, all information disclosed in the documents is seen as information for accountability purposes. The inclusion of outcome information is intended to show that departments/ministers have implemented an outcome-focused operation. However, this convention may not work in practice, since the public may view it as an indication of the departments'/ministries' accountability for outcomes. Consequently, chief executives are seen to be concerned more with the ability to report the information required under the MfO initiative, rather than implementing the real operation of outcome-focused management. A respondent from a central agency argued that:

“... the statements of accountability became the driving force and were not translated into month-to-month management information and approach to running an organisation that reflected a long-term approach. (Respondent 1)

Lastly, when the public sector focuses on outcomes, the inherent difficulties in assessing the achievement of outcomes come into play. The first difficulty relates to tension between the politicians' and the bureaucrats' horizons for their goals and objectives. In New Zealand, politicians face a three-year election cycle, while the leadership cycle in public agencies is longer than three years. Politicians' main concerns are to show short-term results, while the concern of chief executives is argued to be longer-term. Since outcomes can be specified at different levels of generality, it is not easy to decide on the appropriate level of outcomes to be set. If a shorter focus is selected, it may not produce better results in the longer term, and vice versa. The second

difficulty stems from the condition that a myriad of relationships exist between outputs and outcomes. It may take time and a lot of trial and error to find the most appropriate intervention for achieving every specific outcome. If the targeted outcomes are the same over time, eventually the right interventions can be identified and executed. However, the situation is always changing, resulting in a change in outcomes sought, especially if the outcomes are more short-term in their orientation. In this environment, it is highly unlikely that the best mix of interventions can be identified. The last difficulty is that outcomes may not be easily measured. It is difficult to specify performance indicators for impacts expected from the delivery of outputs in some areas of public service delivery. One respondent argued that her department's minister was concerned with the influence of the department's policy on a group of people in society, but it was difficult to find indicators that could show the level of influence of the policy on society.

### **Leadership in the MfO Initiative**

One factor that is cited in some reviews of the MfO initiative, such as the review of the OAG and some other unpublished reviews, is the leadership of the MfO initiative. The first is leadership in the implementation of the initiative. The political support was strong for the initial stage, but this support slowly diminished over time. The initiative was developed by a bottom-up approach, in which central agencies led the initiative and pushed it on to the public management system. There was no parliamentary regulatory umbrella for the initiative. The initiators claimed that the PFA 1989 itself already gave them the statutory reason for the MfO initiative, as stated earlier in the Economics and Strategy Group (2003)'s review. So, there was no need to add another regulation or to modify the current regulation, even though later the PFA 1989 was revised to become the PFA 1989, as amended in 2004. The pressure from the government on agencies was minimal. The pressure from central agencies is not strong enough to compensate for the missing pressure from the government. Leadership was very weak in implementation of the initiative.

The other leadership issue is in the performance reporting area. As the Office of the Auditor General (2008) claims, there is no clear responsibility established for professional leadership and oversight of the preparation of non-financial performance information. There is no professional body that provides leadership for preparing performance information equivalent to that provided in the financial reporting area. There is no standard-setting initiative for performance reporting. No professional body is in place to ensure that there are technical standards laid down for the preparers of

non-financial performance information. Leadership is currently assumed by central agencies and the OAG, but no coherent idea of proper performance reporting is agreed on by these agencies. The preparers are confused because they receive conflicting messages on the quality of, and the need for improvement in, their performance reporting.

### **Declaration of Early Victory**

There is an acknowledgment by the officials of central agencies that the failure of the MfO initiative to implement outcome-focused management and to improve the quality of performance information may be caused by the oversight process and the early declaration of the success of the initiative. Too much focus was on visible things, such as what the reports produced, when the initiative was evaluated formally and informally, once it began to be implemented. There was not much initiative for a comprehensive investigation of the actual management process itself. One respondent from a central agency stated:

... I think in hindsight we put too much emphasis on the SOI and not enough emphasis on the change processes that would support strategic management better. (Respondent 1)

The presence of an outcome framework in most of the SOIs issued during the earlier periods of the MfO implementation has led officials in the central agencies to believe that agencies were able to implement outcome-focused management. Based on the results of early evaluations, they believed that the presence of an outcome framework showed that chief executives understood the outcomes they were trying to achieve and the way the outputs contribute to the outcomes. However, their belief was later to prove untrue. They declared victory too early. Another respondent from a central agency described this situation:

Central agencies shifted their attention to other things too quickly. Once every agency had done at least one annual cycle, we considered that the knowledge was out there and chief executives and their senior team would run it on its merit. I think in hindsight we were wrong. We declared victory too soon. (Respondent 3)

Once they declared victory, the attention of central agencies moved from the MfO initiative to other activities, until they realised that the adoption of outcome-focused management had failed.

## **Level of Support**

Another factor that led to the failure of the MfO initiative to produce the intended impacts was the low level of support from central agencies. There is a little bit of a “blame game” in the claims about this factor. The majority of respondents claimed that the level of support was not enough to help them to implement outcome-focused management. However, respondents from three agencies acknowledged that the level of support was adequate. Two of the three agencies were from the group of the top 9 agencies, based on the results of the disclosure analysis; the other agency is from the group of the middle 9 agencies. Respondents from four agencies claimed that the level of support was not sufficient to enable them to produce good outcome information. The composition of this set of these agencies was: one agency from the top 9 group, one agency from the middle 9 group, and two agencies from the bottom 9 group. The respondents from two agencies, one from the middle 9 group and the other from the bottom 9 group, expressed frustration with the level of support from central agencies.

On the other hand, two respondents from central agencies argued that the support was available but agencies did not show initiative in seeking support. One of these respondents commented on this tendency:

... How many of them even know where the Pathfinder project is? How many of them regularly attend CFO network meetings? How many of them seek out the people who are identified and are still around in the Treasury on financial management issues? Not very many. It is easy to criticise support, it is much harder to sustain that criticism if there is no evidence of any efforts to seek that support. (Respondent 3)

This blaming game may indicate two issues. First, there are different sets of incentives at play for agencies in implementing outcome-focused management. Some chief executives may see the merit of the initiative and seek diligently any available help and guidance to ensure that they can produce a good quality of outcome information. On the other hand, other chief executives may not see the same incentives and instead they may view the requirement for outcome reporting as another compliance exercise. These executives may pressure their staff to come up with something to fulfil the reporting requirement. The staff may not have good access to such support and help. Second, the initiative was not organised in a coordinated way: the support mechanisms were scattered over many places, so that it was not easy for agencies to obtain them.

## **Accountability Shifting**

The last factor that may contribute to the failure of the MfO initiative to provide a good quality of performance information is the concern that most of the accountability required in the public sector has been shifted from ministers to chief executives. The doctrine of ministerial responsibility has been modified into a doctrine that calls for joint responsibility from ministers and bureaucrats. There is a convention that ministerial accountability for departmental action means that the minister should answer to the public for such action and leave the public servant anonymous, regardless of their personal responsibility (Mulgan, 2008). Internally within the government, the minister will take any necessary steps to impose remedies and sanctions. Ministers are the public face of the bureaucracy in the political realm.

Even though the overall structure of ministerial responsibility still remains, the MfO initiative has expanded the accountability of agencies into the area that previously may have been the responsibility of ministers. At least by implication, the accountability is shifted away from ministers. If there is public failure, the public may have the perception that chief executives are to blame. Chief executives are facing greater risk due to the increased accountability which MfO demands of them.

### **4. Reporting Incentives in the New Zealand Public Sector Setting**

The incentive for reporting outcome information is different from the incentive for the established practice of output reporting. In output reporting, ministers were acting as the principal for chief executives in demanding performance information. The interviews revealed that ministers used more detailed performance information in the output agreement to discipline chief executives. There is external pressure, from the perspective of chief executives, to put resources into the information infrastructure. Chief executives used performance information in managing day-to-day operations to meet the expectations of ministers. In addition to this, people understand that output is the responsibility of chief executives. Where there is a concern about bad news from output information, it is highly likely that chief executives, rather than ministers, will be questioned about it. There is therefore an incentive for chief executives to hide this bad news as long as possible, but there is an opposite incentive, from the ministers' perspective, to reveal this information, at least internally between them, since it is in their interest to act on this bad news to ensure that their priorities are still going to be achieved.



The prescription of outcome information has a different incentive setting. Outcomes are still the responsibility of ministers. They give attention only to the outcomes set in their own interest, in terms of priorities that they can control. Similar to chief executives, they want to disclose information on aspects that they can control. The findings indicate that while ministers are concerned with the short- or medium-term impact of their priorities in the horizon of the election cycle, they do not involve themselves in the development of outcomes from service delivery that may not materialise in the short or medium term. Even for information on aspects that they can control, like any other agents, ministers have a preference for releasing good news sooner and withholding bad news as long as they can, because of the asymmetric value of information. As has been stated above, the incentives for ministers and chief executives are aligned in outcome reporting.

### **Incentives in the Budget Allocation Process**

Budget allocation can be used as a strong reporting incentive. The findings reveal that chief executives and ministers will be likely to disclose more information if they are competing for more funding in a period with an increasing budget. Once the funding is given in one year, it will continue to be given in future years, as long as the government budget increases. For ministers, the disclosure of performance information is important in the public debate within the cabinet and select committees in order to get more resources that can be allocated to their area or priorities. In the New Zealand setting, chief executives have greater flexibility in allocating resources. For them, an increase in funding is always favourable in that it is able to produce good news or increase their perquisites. The interests of ministers and chief executives are thus aligned.

On the other hand, these interests are not likely to align in the period of a budget cut. Some respondents stated that there is not much debate when reduced funding is sought. Chief executives will be likely to disclose the sort of information that supports avoiding a funding cut, but ministers want to get another type of information so that they can set how much funding they can cut. The situation is exacerbated, because input information is no longer available in the budget system after input control was transferred to chief executives following public sector reform. Ministers find it difficult to make informed decisions. In one case that occurred outside the period focused on in this study, the current government had to employ former bureaucrats with extensive knowledge of government department operations to find areas in which funding could

be cut, as a consequence of the economic crisis resulting in a decreased government budget. This case highlights an important factor: the problem of bounded rationality is quite marked in the public sector, in which principals do not have enough expertise even to ask for the right information that they need, to deal with their agents.

The findings from the interviews revealed that performance information is not used in the management of the public budget. One interviewee who did a review of accountability documents argued that performance information is always ignored, because the main control on public spending still revolves around input control. He further stated that even though it is almost two decades since public sector reform changed the focus of public sector management from input control to output control, deep within the system people are still controlling inputs instead of outputs. The budget allocation and evaluation discussions in the select committees of parliament are more about inputs. In this setting, nobody in the budget system provides the incentive for output and outcome reporting, since at the top level (parliament) the concern is with inputs.

### **Incentives in Evaluation and Performance Assessment**

Another underdeveloped area in the management process is the closing link in the process, which is performance assessment and evaluation. While so many efforts have been directed to develop better planning, such as the introduction of a strategic management framework and the MfO initiative, not much attention has been paid to the assessment or evaluation of what has been done and whether what has been planned has been achieved. A periodic comprehensive evaluation, such as an impact evaluation, is needed to assess whether what was planned has been delivered efficiently and effectively. In impact evaluation, there is an in-depth analysis of the causes and effects and a search for opportunities for greater value for money from agencies' service delivery. This type of evaluation is rarely performed in the New Zealand public sector. Where such an evaluation does occur, it is more likely that it was done on an ad hoc basis (Ussher & Kibblewhite, 2005).

While some efforts are still needed to improve the planning process in outcome-based management, the focus of the MfO initiative, if it continues, is to develop a culture of evaluation in the public sector. There is some degree of evaluation of new spending, because Treasury and ministers will engage in more scrutiny of this spending; however, there is limited attention to evaluating the existing level of spending. When new policies are proposed, Treasury asks agencies to provide indicators to measure the

impact of the policies. Unfortunately, the share of the budget for new policy funding is very small. A large amount of the budget is allocated to existing or current policies that, while they may still be important policies, may not be scrutinised rigorously by central agencies. If external evaluation of these policies is not feasible in the short term, it may be useful for the MfO initiative to make it a requirement that agencies show that they are evaluating their existing policies. The disclosure that they periodically make of the appropriateness of outputs to achieve outcomes is as important as the disclosure of outcomes. Alternatively, a formal external evaluation of the performance of public sector agencies can be set up. In other jurisdictions, an independent agency is established to assess the performance of other public sector agencies. If no mechanism can be established within the current central agencies, there should be a new agency established to perform this function. It seems that the recent establishment of the New Zealand Productivity Commission is intended to fill part of this void.

Performance assessment can be used to induce chief executives to disclose more performance information. Interviewees' attitudes towards performance assessment were quite favourable, especially if the assessments were based on output information. To some degree, there was some support for performance-based compensation. However, the current performance assessment system at chief executive level is quite complex, because there are so many factors included in assessments. Achieving performance, as reported in the accountability documents, is just one item in a list of many outlined in performance assessment forms. Therefore, it is highly unlikely that the current assessment system has any impact on performance reporting practice.

In order to put stronger pressure on chief executives to improve the quality of their performance information, their performance assessments should provide a more balanced focus that includes an evaluation of non-financial performance targets. Norman (2002) argued that while there are severe consequences for failing to meet financial performance requirements, there is not an equal penalty for failing to deliver on quality and outputs. Chief executives will not give enough attention to providing a better specification of outputs unless they are held responsible on under-delivery for these areas in the same way as they are judged if they miss financial targets. Similarly, chief executives will be induced to produce better outcome information or implement outcome-based management if they are evaluated by their ability to specify and monitor their outcomes and to perform systematic evaluation of how well their outputs contribute to outcomes.

## **Incentives in the Information Control Process**

The strongest incentive for improving the quality of information is the presence of information quality control. In the current arrangement, auditors are responsible for verifying the information reported. In the New Zealand setting, their position is independent and they work for parliament. Their role, which is mandated by the PFA, is limited to auditing financial information and output information. Auditors were consistently cited in the interviews as the only party concerned with performance information reported in the accountability documents. Auditors use their leverage in issuing the audit opinion, as well as recommendations in management letters, to pressure chief executives to improve the quality of information over time. Auditors are able to ensure the reliability of information, since they verify the information system that supports chief executives' audited assertions. It was found that longitudinal consistency is quite strong in output reporting.

Despite the strong role of auditors in controlling the quality of information, the current external quality control does not elevate information quality to a level that is satisfactory, even to the auditors themselves. A high degree of cross-sectional inconsistency is found in output information. Two factors contribute to this deficiency. First, there is a high degree of discretion seen in the reporting. Second, auditors had not yet focused their audit on the relevance of information. In terms of outcome information, since it is not audited, the auditors' pressure is viewed as not particularly effective in improving quality. The lack of authority in verifying the information, coupled with some degree of disagreement on how to report outcome information, has meant the quality of outcome information has improved very slowly, which is very disappointing. The Economic and Strategy Group (2003) has raised the need for a mandatory quality assurance process in the MfO initiative, including an external review of the content of the SOI, after performing an early review of the MfO initiative and the Pathfinder project. Unfortunately, this suggestion was not followed up by central agencies, resulting in the current situation of disappointing quality in performance reporting.

## **5. Challenges in New Zealand Public Sector Performance Reporting**

Several challenges are identified in this study in relation to performance reporting practice in the public sector, especially in the New Zealand setting. The first three challenges relate to the external factors affecting the reporting entity: the

principal's bounded rationality, information quality control and reporting policy advice outputs. The next three challenges are internal ones for the reporting entity: complexities in managing for outcomes, determination of the appropriate level of outcomes and the current state of information technology.

### **Ministers' Capacity and Capability**

Quality of information will improve if the problem of the principal's bounded rationality can be mitigated. If the principal knows what kind of information to demand from the agent and the principal is able to understand the agent's information and use it for decision-making in his/her relationship with the agent, then better information will be reported. The need for this capability is not cogent in the case where the information incentive is the same for ministers and chief executives. Usually, during a period of prosperity, ministers largely ignore published performance information, so no pressure is directed toward chief executives to improve the quality of information. Only when the incentive is not aligned, such as in the case of a funding cut, is the principal's capability in relation to information acquisition and processing highly valued.

The capacity and capability of ministers has been considered a serious issue since the original New Zealand model was implemented. A review by the Steering Group (1991), as quoted by Boston (1992), described a concern as to whether ministers had adequate support, in the form of independent advice, when developing and negotiating performance agreements with their chief executives. In the review, ministers were portrayed as extremely busy and many of them did not possess technical qualifications in the policy area of their portfolio. Consequently, they also did not have time or the technical knowledge to actively participate in the specification of their department's performance, especially in framing the desired outcomes, selecting appropriate departmental outputs and setting proper performance standards.

This capacity and capability concern is becoming more complicated due to the severe asymmetric situation, as previously explained, in which agencies hold virtually all the information needed to assess current performance and future capacity. Ministers must be heavily dependent on the advice of chief executives and other senior officials to decide the specification of outputs that ministers (government) have to purchase from departments/ ministries. The interview results show that ministers rely on their chief executives to enable them to make decisions on output plans. The typical process for decision-making on outputs is that ministers give their directions at the outset and ask

their chief executives to translate them into a set of outputs which the ministers will consider later and decide whether to accept or revise.

The interview results also showed that while some ministers provide a sense of direction on what their departments have to do in the next year, they do not actively participate in the planning of the long-term strategic performance of their portfolio of agencies. The process is mostly one-way. There is no structured debate on strategic direction, capability and performance so that agencies can properly report their ability and performance to parliament. It depends on chief executives and agencies to come up with their own long-term strategic vision and the strategies to retain long-term capability. This tendency is quite marked in agencies that provide service delivery instead of policy advice. There seems to be a combination of factors leading to the lack of an incentive for ministers to come up with long-term strategic plans, partly due to the adversarial environment of politics in the Westminster system, or to the lack of capability on the part of ministers to provide more than a sense of direction to their chief executives.

There is a need to provide greater support for ministers so they are able to perform their roles as active and discriminating purchasers of departmental outputs. Such purchasers should demand that departments/ministers produce clear expectations about service delivery and innovatively propose several sets of outputs to meet these expectations. Ministers should be in an environment in which several choices are available, instead of the current process in which agencies come up with the set of outputs that they think is the most appropriate. Ministers should be able to select from several options (each with its consequences outlined) that are put on the table by their departments.

Several suggestions have been offered to mitigate the problem of the bounded rationality of ministers, as stated by Boston (1992). The first suggestion is that ministers should be given more resources, in the form of additional staff in their office who are specifically assigned to assist them with the development and negotiation of the output plan. This suggestion seems to be the most promising one, considering that it addresses the problem directly by the provision of support that gives ministers the ability to have more control. However, the possibility of employing staff with the appropriate level of expertise and experience may be small unless a considerable amount of funding is allocated to the ministry's office. If the amount is too high, it may attract public concern about the appropriateness of the spending.

The second is that the SSC or another central agency is given a mandate to help ministers in dealing with their departments in matters related to the planning and evaluation of performance. Currently, the central agencies' responsibilities in this area are unclear and unspecified (Cook, 2004). If they are given a clear mandate for this role, more resources need to be allocated to central agencies to enable them to perform this role. The limitations in the first suggestion, which is the availability of suitably qualified officials assigned to perform the tasks and the funding, also apply to this second suggestion.

The third suggestion, which is the strategy that has been adopted in the past and is considered to be the right strategy for the future, is employing a number of high-quality purchase advisers. The practice of employing advisers to ministers to provide a different source of policy advice has become more popular since the MMP voting system was implemented (Eichbaum & Shaw, 2007). To be able to provide meaningful assistance, these advisers should be senior managers from the private and public sectors. They should have a level of expertise and professionalism that matches the expertise and experience of the current bureaucrats. Scott (2001) claimed that ministers who employ purchase advisers tend to be able to perform as discriminating purchasers. There is a tendency for the quality of output specification to be higher in agencies whose ministers utilise purchase advisers. In the past, these advisers were employed ad hoc by the ministers of finance. Since they are not a part of the formal and regular system, there is concern about their accountability and their cost-effectiveness. Recently, the New Zealand government took a formal approach to providing advice to ministers related to the strategic planning, performance and capability of their agencies by establishing the Productivity Commission.

### **Information Quality Control**

Some interviewees argued that a lack of meaningful performance information can be alleviated if there is an information quality control function located internally in the government itself, in addition to the role of the OAG in providing assurance externally. In other jurisdictions, there is a dedicated unit responsible for analysing performance information which can ask agencies for more information in relation to their functions. The employment of purchase advisers that was previously mentioned highlights the need for this internal quality control. One respondent argued that this role would be appropriately included in Treasury, while another respondent from a central agency argued that the establishment of the Productivity Commissions (outside the

observation period of this study) might fill this gap. These suggestions highlight the need for internal quality control of performance information if the improvement in information quality being sought is to be achieved.

A lack of reporting standards has contributed to a great deal of confusion and frustration among agencies in relation to the disclosure of their performance information. It also results in cross-sectional inconsistency in reporting. The findings show that the preparers of performance information felt they were unfairly treated. On one hand, they were criticised for not reporting properly, while on the other hand they did not know how to report properly. The availability of reporting standards would increase the likelihood of disclosing relevant performance information. A similar note was also raised by the Economics and Strategy Group (2003) in their evaluation. There should be improved guidance available and wider distribution of a good practice model, to support agencies in developing performance information. The relevance characteristic is still not properly addressed in the current arrangement, as auditors concentrate more on the reliability of information and nobody in the system checks the relevance of information, except the reporting entity itself. However, developing reporting standards is not an easy job. Currently, it sits with the accounting profession but their main focus has been on financial reporting. Improvement in this regard may take place with the establishment of the External Reporting Board in 2011 and if the Board elects to set separate standards for the public sector.

### **Reporting Policy Advice Outputs**

Interview results showed that the biggest challenge in output reporting is the development of a meaningful set of output measures on policy advice. The current practice of measuring performance areas of outputs as if they are just a document does not take into account any assessment of the achievement of delivering policy advice to ministers. On the other hand, an alternative practice of considering policy advice activities as part of a bigger output of policy creates a concern as to how to link advice or a series of pieces of advice to the successful application of a policy. Because the links are not so obvious that does not mean that the quality of the policy advice should not be assessed. However, a good policy may not be accepted because it is not the policy that is the preference of the political principals. Equally, a bad policy may be accepted because it meets the preference of political principals.

The right measures are the ones that drive public officials in managing the delivery. In the case of policy advice, the proper measures are the ones used by chief



executives in the delivery of policy advice to ministers. The key function of policy advice is to inform ministers about any problems identified and to propose interventions to tackle the problems. This involves an identification of the impacts/outcomes needed to solve the problems, several choices of outputs that can be delivered to achieve the intended impacts, and a justification of the link between outputs and outcomes based on evidence and sound reasoning. What is essentially produced by policy advice activities is a complete set of outcomes framework and measures of an intervention or a series of interventions for the proposed policy. Policy advice outputs should be presented in a proper context, which is to provide justification for an intervention and to provide mechanisms to assess the performance of the intervention. When performance measures of policy advice outputs are developed in this way, they serve two purposes. First, the measures developed are consistent in the way they deliver policy advice outputs. Second, it makes the appropriate and comprehensive outcomes framework available for all of a department's/ministry's interventions.

### **Complexities in Managing for Outcomes**

The third challenge is inherently related to the architecture of the MfO initiative. This initiative tries to develop a system where chief executives are accountable for output delivery, responsible for providing outcome information to enable the ministers to select the appropriate outputs and at the same time effectively manage for outcomes. This initiative requires a clarification of the distinction between non-accountability for outcomes and managing for outcomes. It needs the type of information that can be used to evaluate and to assess whether chief executives effectively manage the outcomes. What is the difference between this information and that which is used in accountability for outcomes? It is not an easy task to differentiate between the two. The impetus to provide a clear distinction will be stronger when the audit of performance information and assessment of performance becomes more formal.

When there is no clear and fully articulated model of the relationship between outputs and outcomes, management will find it hard to defend their choice of outputs to achieve the targeted outcomes. In this situation, the most difficult aspect of implementing the MfO initiative is the requirement to develop credible intervention logic and an effective structure for performance measurement. The interview results showed that even though agencies could produce a skeletal form of outcomes framework, it was hard for them to fill the empty spaces with explanations to justify their framework and sets of integrated measures for all components in the framework.

What is needed is the provision of extensive guidance or a comprehensive expenditure management model of outcome-based management that can be used by all agencies to develop justification of the critical linkages.

Managing effectively requires a system in which agencies have a chance to experiment and to learn from mistakes in identifying the appropriate mix of outputs to achieve the targeted outcomes. But, considering that there is an asymmetrical value of information in the political setting, there is not much room to make mistakes. Nobody can guarantee that the potential principal will not take advantage of an agency's mistake to score a point. Here, the challenge is to develop a performance evaluation and assessment system that can inform whether chief executives are held accountable for output delivery and are moving in the right direction in effectively managing for outcomes, while at the same time allowing them to experiment to find better ways of achieving outcomes without promoting negative incentives on behalf of the current principals. The likelihood of developing such a system is very small.

Developing a system in which the ability to learn from mistakes is fostered is the greatest challenge faced by agencies. Nobody will experiment to learn from their mistakes when others punish them for making mistakes. It requires a change from a culture of quickly blaming someone for a problem to giving them a chance to show that they can do it better next time. The latter is certainly not the culture in New Zealand society and the political setting, regarding problems in the public sector. The Westminster system has made the political setting too adversarial to foster this culture.

### **Determination of the Level of Outcomes**

The fourth challenge relates to selection of appropriate levels of outcomes that chief executives are going to operate. There is a strong tendency to select high-level or long-term and ambitious outcome targets because of the perception that these broader society outcomes are the appropriate ones for the public sector agency, to show that they reach out to society. This tendency is further driven by the public setting that they operate in, which requires them to broaden their reach instead of focusing it. It is easy for them to get an agreement on the outcomes targeted when they are broad; besides, it does not invite unnecessary public debate about the chief executives' performance. Auditors will not criticise them in the management letter, as has been asserted by some interviewees. A high-level outcome framework may inhibit the development of effective and realistic interventions and plans, and therefore it results in the difficulty of linking the outputs produced to the outcomes sought. The shorter term focus in the

outcomes framework increases the controllability or influence of the agency on outcomes, and thus makes it easier to understand the intervention logic used and develop effective plans. However, the incentive for a shorter focus is much smaller than that for a longer focus.

Another advantage of focusing outcome measures on short-to-medium term outcomes, such as a 2-5 year window, is that they will be aligned with the political cycle of the ministers. A typical minister in a new government has a three-year span in his/her term. The most time that this minister is able to manage his/her portfolio is probably only two years, because of having to spend the first year gaining an understanding of it. In the first term, the minister will probably focus on a two-year window of results, but with a small amount of attention on the five-year window. If his/her coalition gets re-elected to govern for another term, he/she can concentrate on the previous five-year window and update the outcomes for the next five-year window. However, this short-to-medium outcome focus requires a more explicit arrangement to be made when establishing the responsibility of ministers for achieving those outcomes. Reporting should go hand-in-hand with accountability. The accountability function of ministers is missing in this area. Together, reporting and accountability can provide a strong incentive to ministers to choose more realistic and quantitative outcome targets.

The challenge of selecting and maintaining an appropriate level of outcomes is not unique to the New Zealand setting. Most countries that have developed an outcome-based framework in their public management system also face a similar challenge. Webber (2004) found that these countries confront challenges in maintaining a consistent level of outcome definition. In most cases, policies do not produce the intended impacts and there are pressures for the policy design to respond quickly and appropriately to the changing situation. Ministries/departments may be reluctant to change their outcome statements in a system with a strong emphasis on outcomes. As a result, policy design is not responsive enough to provide different directions and consequently different policies. Countries that implement outcome-based management also confront challenges in maintaining the feasibility and practicality of outcome targets. Usually, politicians aim for higher but impractical targets; others, including the bureaucracy, may want something that is applicable and feasible. The challenge is to identify and maintain the appropriate horizon for outcome targets, so that policy impacts can be measured and policy assessment can occur meaningfully. If the horizon is too short, the policies are not stable enough to produce the intended impacts. On the other hand, if the horizon is too long, it may be too late to change a failed policy.

## **The Information System Issue**

The last challenge to be faced is the current capability of the agency to produce performance information, especially outcome information. An early review of the MfO initiative revealed that the task of developing a proper performance report, in order for chief executives to be able to manage for outcomes, required skills and resources that the agencies did not have (Economics and Strategy Group, 2003). If they are pressured to produce such information without proper support from central agencies, the agencies' workload may be distorted. Agencies will not be likely to pursue this direction. A majority of the respondents in this study stated that the priority of agencies is for service delivery, rather than for system information supporting the accountability documents.

The state of the information system in public sector agencies is not as advanced as comparable entities, in terms of size, that are operating in the private sector. One of the examples that was frequently cited by a respondent in this study, in discussing the current state of the public sector information system, is the inability of the system to generate cost information at an individual output level. As service providers, where a considerable number of production costs are in terms of labour costs and overheads, government agencies do not have cost accounting systems to allocate these costs to their outputs. Accessing information that in theory is supposed to be able to be collected from inside the organisations is still very difficult for most government agencies to provide. Pressure to invest in a cost accounting system may come from two sources. First, central agencies, working together with professionals in the accounting field, may develop output pricing methodologies that are suitable for the public sector situation. These will serve as standards on how to price and eventually to cost, since most public services are not contestable, so that proper pricing methods would be some variant of cost-based pricing. Second, audit should direct greater attention to output costing. This strategy will provide a strong inducement for agencies to invest more in their information systems in order to satisfy audit requirements.

Another issue that is worth discussing is where the ability to produce some outcome information is extremely difficult, especially on high-level outcomes. For example, one interviewee discussed the need for fish stock information, since his agency decided that the level of fish stocks was the measure to use for their outcomes. To get this information is costly and requires them to employ fisheries scientists that are in limited supply worldwide. It takes around one-third of the budget spent by a typical fisheries agency in OECD countries to gather information at the current level needed for managing their business. Gathering more information to successfully manage for

outcomes is not viable for such agencies at the present time. Other research also highlighted the same issue. Ussher and Kibblewhite (2005) found that while some departments have developed a comprehensive set of world indicators, they still cannot produce effectiveness indicators and risk indicators<sup>4</sup>.

---

<sup>4</sup> Ussher and Kibblewhite (2005) define each of these indicators as follows:

- World indicators provide a measure of the targeted world or aspect of it. For example, the indicator for a children's education service is the percentage of children sitting and achieving School Certificate in 5 subjects
- Effectiveness indicators show the measure of the success of particular interventions, such as the reoffending rate of a certain group of offenders that receive a particular intervention, compared with a control group.
- Risk indicators provide information on the area to which the intervention should be targeted, for example the risk of imprisonment within 5 years.

## **VI. SUMMARY, CONCLUSION &**

### **RECOMMENDATIONS FOR FUTURE RESEARCH**

This chapter presents the conclusions and recommendations of this research study. A brief summary of the study is presented first, before the conclusions are discussed. The chapter concludes with recommendations for future research.

#### **Summary of the Purpose of the Study and the Research Problem**

The aim of the study was to fill the void in public sector research relating to performance reporting practice in New Zealand. First, as the output-based New Zealand model has morphed into a balanced focus of output and outcome model of the public management system, there is limited research directed to address the impact of the new focus on reported performance information. In addition, with just a few exceptions, most of the research and reviews have concentrated on the early process of change. This study was intended to provide a picture of the long-term change in performance reporting practice of the newly modified New Zealand model. Second, academic review of the MfO initiative to incorporate outcome information into the accountability documents has been limited. The few reviews that have taken place have used official documentation as the source of data. This study aimed to provide more evidence, by securing comprehensive inputs from various actors in the public sector regarding the issues found in public sector reporting. Third, some results from external assessment of the quality of performance information were available, but they were not in a detailed format in which the results of individual assessment of each agency were presented. In addition, there was no discussion on how the assessment was conducted, including the benchmark and methodology used. Studies focusing on non-financial performance information have mostly measured the quality of information using a simple methodology of indicator counting and categorisation. This study aimed to provide a more sophisticated methodology to measure the quality of non-financial performance information disclosed in the accountability documents of public agencies.

With these objectives in mind, the study focused on measuring the quality or comprehensiveness of output and outcome information and on identifying the incentives and challenges in the practice of reporting this information. Specifically, the study aimed to address a series of questions related to this focus. First, was there any justification for the assertions made by the OAG about the poor quality of performance

information? Second, how was the quality of performance information measured? Third, what was the impact of the MfO initiative on performance reporting practice? Fourth, has there been any improvement in the reporting practice of outcome information? Fifth, how useful was the reporting guidance or benchmarks? Sixth, what were the factors that affected reporting practice? And last, how did various actors interact in performance reporting practice?

### **Summary of the Research Background**

The current New Zealand public management system is an evolution of the new model introduced in the late 1980s. The model included a clear distinction being made between the responsibilities carried by ministers and those handled by chief executives in the operation of the government and its bureaucracy. The model was based on the implementation of some specific economic theories, particularly public choice theory and principal agent theory. The aim of the model was to enable the public sector system in New Zealand to be more efficient and effective in delivering its services.

In the model, a specific principal-agent relationship was established between ministers and their chief executives. Ministers were designated as the principal who had the ownership and purchase interest in the department. In performing their role as purchasers, ministers were responsible for outcomes. Chief executives were the agents employed to manage the ministers' ownership and deliver on their purchase of public services. A contract, in the form of an output plan (purchase agreement), was agreed to by ministers and their chief executives in the planning stages. The contract served as a basis for chief executives to deliver public service. At the end of the period an assessment was made, requiring chief executives to provide information on the performance of their delivery in the form of a Statement of Service Performance that contained detailed information on outputs planned and delivered.

While the focus on output was established, there was no visible mechanism put in place to ensure that the New Zealand public management system had an equal focus on outcomes. There was a concern that the selection of outputs did not include outcomes for consideration. To mitigate this concern, two attempts were made to include an outcome focus in the system. The first attempt was based on the strategic management framework used in the private sector. The second attempt, initiated after the failure of the first one, was intended to develop a more suitable approach to outcome-based management in the public sector.

The first attempt failed mainly due to the incompatibility of the framework with the public sector setting it was adopted into. The strategic management framework was a combination of two unrelated initiatives aimed at modifying the original New Zealand model that was considered to be flawed. The first initiative was a top-down approach, in which the government of the day provided a set of directions for bureaucrats in the form of the SRAs framework. At the same time, the bottom-up process of identifying the best indicators to assess the performance of chief executives was initiated within central agencies. To bridge these two initiatives, a set of indicators called KRAs was introduced. The KRAs were later deemed to be incompatible with the existing output performance reported on in the Statement of Service Performance. A change of government spelled the end of this initiative, and it later largely abandoned.

The second attempt, called the MfO initiative, is currently being implemented. The new system puts additional responsibility on chief executives to show that they are managing their output delivery to achieve certain targeted outcomes, even though they are still not responsible for outcomes. This initiative has resulted in the disclosure of a new set of information on outcomes, to complement the existing reported information on outputs. This study has been done in this new setting, and has focused on reporting practice of these two types of information.

### **Summary of Theoretical Framework of the Study**

The study used public choice theory and agency theory as the framework for analysis. The rationale for selection of these theories was that they are the theoretical framework used as the foundation for the reforms.

A public choice theory framework was used to identify public choice problems and their causes. The problems included policy capture by interest groups or specific constituents and bureaucratic capture. Policy capture happened because a politician's authority over allocating rents in the public sector was not controlled. On the other hand, bureaucratic capture was mainly the political principal's problem of controlling their executive agents, which are agency problems in a typical principal and agent relationship.

Public choice theory offered some strategies to help mitigate public choice problems. These strategies were based on the idea that the intervention of government should be minimal and instead closer attention should be given to how individuals interact with each other. Three of these strategies were considered within the theoretical framework. First, there should be an effort to limit a politician's authority, so that the



total allocated rents can be reduced. Second, there should be a separation between policy-making and policy delivery activities, so that bureaucratic capture could be minimised. The policy-making agencies should be induced to become a proxy for the consumers' voice in the public sector, to inform the policy delivery agencies of the mix of outputs desired by the consumers. Third, proper incentives should be installed in the public management system so that public agencies would deliver socially desirable outputs and outcomes. With these incentives, the self-interest of public officials could be aligned with the interests of society regarding public service delivery.

Information disclosure played an important role when consideration was given to each of these three strategies. Disclosure of output and outcome information enables voters to assess the adherence of politicians to the limitations imposed on them. This information would also be useful for consumers of public sector services to assess whether government agencies have addressed their concerns. Lastly, output information could be used to assess whether public officials delivered socially desired outputs.

Agency theory was used to identify factors that might support or inhibit performance reporting practice. Specific aspects of the theory were selected, those that were mainly present in the public sector setting. These particular aspects were: verifiability or measurability of outcome/performance, ex-ante causal ambiguity, the presence of multiple principals, severe bounded-rationality on the part of principals, a high degree of risk-averse behaviour on the part of agents and the level of goal conflict in the relationship. Two additional mechanisms were also included, which were the role of reputation in the agency relationship and the presence of severe asymmetric information.

### **Summary of the Research Methodology**

Two research methods were used in this research. The first method was the disclosure index, used to address concerns relating to the disclosure and quality of performance information. The second was the case study, which was used to address research questions relating to issues concerning performance reporting practice.

The disclosure index was used after considering some alternative ways of measuring information quality. This decision was also an attempt to answer the research question of how to measure information quality. There was no direct measure of quality of information available, since quality itself is hard to define. Most external proxies assume that perception of quality is the measure of quality, while internal proxies assume that the extent or comprehensiveness of information is the measure of

information quality. Since external proxies were not available, the most appropriate approach to use was the disclosure index technique, which is the most popular and most advanced technique for measuring narratives in a report. It is flexible enough to incorporate qualitative and quantitative aspects of the information. The process of development, and later the application, of the disclosure index was carefully implemented to ensure that the methodology had a high degree of reliability and validity. The validity of the index was strengthened by ensuring that the index was validated by experts and experienced practitioners in the New Zealand public sector management system. The reliability of the index was tested by ensuring that the difference between scores in the test-retest procedure, as well as the inter-coder reliability test procedure, was less than 1%.

Multiple case studies were selected, due to the need to gain richer data to enable the identification of various concepts and aspects discussed in the theoretical framework. Nine agencies with the role of preparers were selected for the study. These agencies were randomly selected, based on the result of the disclosure analysis. In addition, two institutions in the central agencies – the Treasury and SSC – and the OAG were also selected, to provide corroboration and possibly different perspectives on the same issues. Interviews were the main method of data collection. A document study was used as a secondary method, to complement the results gathered from interviews. The results from interviews and the document study were coded and analysed based on the theoretical framework used.

## **Conclusion and Impacts**

The results of the disclosure analysis show that the quality of performance information is still low. This is consistent with the assertions made by the OAG on the same matter. In terms of output information, the level of information provided was, as predicted, stable over the period of analysis, as the practice of disclosing this information had been established for more than a decade before the period of analysis. The disclosed information for the whole period of analysis was around 69% of a complete set of information, as an average rate. The only measure of output information that was disclosed comprehensively was the quantity measure. The quality, timeliness, location and cost of outputs were only partially disclosed. In terms of outcome information, there was a large and noticeable initial jump in the level of outcome information provided, but a very small increase of information over time. On average, only around a third of a complete set of outcome information was reported by agencies

six years after the roll out of the MfO initiative. The quality of disclosure is good in only two areas: the outcome definition and the outcome framework. Other aspects, such as outcome measures, intervention logic (attribution from output to outcome) and risk identification, were only in part disclosed or just poorly disclosed.

There is no indication that output information has improved because of the MfO initiative. There is also no indication that agencies made adjustments to their outputs in order to align them with the outcomes selected. There is considerable evidence that agencies tried to give information in the accountability documents that showed they identified their outcomes and the link between their outputs to various levels of outcomes. However, there is no indication that these agencies are continuously monitoring outcomes, since they were not able to provide targeted indicators for their outcomes. The only indicators that some of these agencies could provide were actual indicators of outcomes, but there is no evidence that links the achievement of outcomes with the outputs that they delivered.

While it must be acceptable for the quality of outcome information to have been very low in the earlier years of the MfO initiative, it is disappointing to see that the quality, as measured by comprehensiveness of information, did not increase rapidly in the later years. There was only gradual increase in the comprehensiveness of information supplied over time. There were some noticeable jumps in the comprehensiveness of information provided in the second year for three agencies, indicating that they had learnt from their experience in the first year. One agency showed a decreasing level in the quality of outcome information being provided over time. Overall, it seems to be quite a difficult challenge for agencies to come up with credible performance information.

The score results may indicate that compliance is the main motive for agencies for performance reporting. Professional guidance (TPA-9), which is not authoritative, was largely ignored, resulting in some mandatory information not being disclosed in the accountability documents. This information, such as an explanation for output class changes, variance disclosure and cost allocation disclosure, is either barely reported or not reported at all. Consequently, some recommended and voluntary information is also barely disclosed. Most of the respondents claimed that they followed Treasury guidance in the preparation of the SOI and the annual report. The disclosure analysis results support this claim for some components of information. It seems that agencies followed the guidance given to the minimum level required. The first thing that they did was to use the same format. Then, they filled it in with information that they could produce.

The indication is that some targeted and actual measures of output were not reported in accountability documents. Furthermore, most indicators included in outcome information were not presented in quantitative terms. Some of the information required by Treasury guidance was also not reported, such as measures of and an explanation for cost-effectiveness and organisational health and capability.

Three factors act to promote improvement in the quality of information provided by public sector agencies in the New Zealand setting. The first, and most important factor, is audit pressure on agencies. The OAG consistently and periodically pressured chief executives to provide better information over time. Another strong factor is the reputation of senior officials in the agencies. The concern about reputation is seen as the motivation for ministries/departments to provide more information, rather than just meet compliance requirements. Those chief executives who were also known as pioneers in applying outcome-based management to the New Zealand setting were able to produce more comprehensive outcome information in the earlier years of the MfO initiative. The third factor that promotes more disclosure of performance information is the presence of a possible funding increase. To compete and eventually secure an increase, agencies were willing to provide more information.

The analysis of data gathered from the case studies identified several factors that contributed to the lack of meaningful progress seen in New Zealand performance reporting practice. First, the MfO initiative has changed the reporting incentives of ministers and agencies from the one that existed under the pure output reporting regime. In an output-based system, ministers may have an incentive to demand more information on outputs, but for outcome information, ministers seem to be less interested. They are responsible for outcomes, so they are less inclined to make this information available. Second, no incentives were in place to induce agencies to disclose more performance information. Performance information was not used at all in budget allocation mechanisms and only in a very limited way in the performance assessment of the agency chief executives. In addition, the lack of authoritative reporting standards in performance disclosure has limited the role of auditors in increasing the quality of information.

The main actors in outcome-based management are ministers and chief executives. They should work together to set outcomes and later decide the most appropriate mix of outputs to be able to implement outcome-focused management. However, the only actor that is the focus of the MfO initiative is the chief executive. The other supporting actors, the central agencies and the OAG, cannot come up with the

leadership necessary to ensure that the quality of information produced by public agencies improves over time. All the actors, in following their own interests, contributed to the current disappointing level of information being provided. Ministers are concerned with their own priorities and that no bad news emerges from their agencies. The chief executives' main motive is to comply minimally with the requirements. Central agencies have claimed that they had already done their job of introducing outcome-based management and requiring the disclosure of performance information. The Audit Office stated that their priority was to perform their mandated role of auditing financial information and output information. Nobody was concerned with whether outcome-based management was in operation or whether the outcome information really reflected a public sector operation focused on outcomes.

Lastly, there are several challenges that need to be addressed if outcome-based performance management and accountability is to be progressed. First, there is a need to address ministers' capacity and capability to understand their departments' business, performance and capability. Second, there is a need to develop a culture of evaluation in the public sector, so that periodic reviews can be conducted to assess the impact of policies and the realignment of outputs to better achieve intended impacts. Third, meaningful reporting of policy advice is the most difficult task in performance reporting. Unless the measures employed are ones that drive the delivery of such outputs, the benefits of reporting performance on policy advice will never outweigh the costs. Fourth, managing for outcomes requires an environment of trial and error. In the current adversarial system of politics, it is hard to convince chief executives to experiment to find better ways of achieving outcomes. Fifth, when all actors cannot agree on what level of outcomes should be in operation, outcome-based management will never happen. There is always a tension between selecting a more volatile but visible short-term set of outcomes and a more stable but less visible long-term set of outcomes. Sixth, there is not much investment in the public sector information system. The skills and capability required to implement outcome-focused management is very limited in all public agencies. The information system in the public sector is not sophisticated enough to perform even the simple task of producing information on output costs. It may take a lot of investment to enable the system to monitor outcomes. If the principals desire implementation of outcome-based management, then they should allocate or increase funding to upgrade the information system. In the current setting, in which a strong desire for implementation is not present, there is limited for support spending money on such a system. Lastly, there is a possibility that outcome-based

performance management is just too difficult to be implemented and it is doomed to fail, as some respondents have stated. The disappointed level of the comprehensiveness of outcome information and the difficulty raised by some respondents in their effort to set the target on outcomes or to come up with outcome measures may show that the links between output-based agency performance and complex outcomes is very complicated. In this environment, an outcome-focused performance management is just impossible to be implemented in a structured system.

### **Recommendations for Future Research**

The issue explored in this research is the quality of disclosure. Quality is determined by the purpose of disclosure and for whom disclosure is intended. While measuring quality in relation to the fitness for purpose may be able to be achieved satisfactorily, an attempt to measure quality from the perspective of the readers of any disclosure is very challenging. It requires an assessment of the characteristics of readers. Different readers have different characteristics, making it difficult to determine what is considered to be the level quality that most readers would agree on as being good quality. A lesser challenge, but still a difficult one, is determining quality based on fitness for purpose, which is the challenge addressed in this study. One question regarding this challenge is what should be the content of disclosures on performance that represents fulfilment of the accountability requirement? There is no easy answer to this question.

Several proxies have been offered as measures of information quality in the academic studies but no dominant method has emerged from these studies. The only increasing trend seen was that more and more research on accounting disclosure has used some form of disclosure index methodology. This methodology is increasingly used because it is more flexible and it can be applied to any kind of complex disclosure on specific and general issues.

The present study is the first attempt to develop a disclosure index for non-financial performance information. Therefore, the index is considered an experimental index, in which there is a possibility that the index cannot capture what is intended, which is to measure the quality of output and outcome reporting for accountability purposes. The items used were based on the requirements indicated and suggestions arising from official documents and literature. The required and suggested items of information used may be appropriate in a small and narrow context, in this case the context of New Zealand central government agencies during the implementation of the

Managing for Outcome initiative; but they may not be appropriate when the context changes.

A similar study could be undertaken on a large scale. One way to do that is by developing the index in the context of other types of government institutions, such as Crown entities or local government agencies. This may lend more credence to the index methodology used. A larger context may result in a broader application of the measure, resulting in the development of a more general index able to be used in assessing the quality of performance information produced. A more general index could have a broader appeal when it is being used in other studies investigating the quality of performance information in the broader New Zealand public sector.

Similar studies could also be undertaken in other jurisdictions. The limitation of the current index is that it is based on the setting of the New Zealand public management system. A more robust index could be developed if some comparative studies were done of countries adopting a similar focus of output and outcome in their public management system. The index developed in these studies could then have a much broader appeal and would become a universal index for measuring non-financial performance information. As more research uses the index, the construct validity of this measure will improve.

This study is also exploratory in nature, in its effort to investigate performance reporting practice using multiple case studies. The findings are preliminary and require further investigation of many of their aspects and dimensions before generalisations can be drawn. The depth of the study is compromised by the constraints encountered. Case studies that are more comprehensive are needed to further clarify details of the findings in this study. Furthermore, this present study targeted preparers and external quality controls on performance information reporting. A study that targets more actors in the public sector arena, such as the inclusion of ministers and members of select committees, may provide additional depth to the findings,

Finally, this study investigated the long-term process involved in the adoption of the MfO initiative. The findings are constrained by the theoretical framework used. If the findings were analysed within the context of a series of historical events they might provide a different but complementary view of the adoption of the initiative. A historical study of the MfO development and its adoption, targeting the main events and actors in the initiative, could usefully place the findings reported in a study of this kind in a proper historical context.

## REFERENCE

- Adams, C. A., Hill, W.-Y. & Roberts, C. B. (1998). Corporate Social Reporting Practices in Western Europe: Legitimizing Corporate Behaviour? *British Accounting Review*, 30, 1-21.
- Adams, M. & Hossain, M. (1998). Managerial Discretion and Voluntary Disclosure: Empirical Evidence from the New Zealand Life Insurance Industry. *Journal of Accounting and Public Policy*, 17, 245-281.
- Advisory Group on the Review of the Centre. (2001). Report of the Advisory Group on the Review of the Centre. Wellington: State Services Commission.
- Anderson, B. & Dovey, L. (2003). Whither Accountability? Wellington: State Services Commission.
- Anthony, R. M. (1978). *FASB Research Report, Financial Accounting in Non-Business Organisations*. Stamford, CT: FASB.
- Audit Office. (1999). The Accountability of Executive Government to Parliament, in Report of the Controller and Auditor-General: Third Report for 1999. Wellington.
- Baker, G. P., Gibbons, R. & Murphy, K. J. (2002). Relational Contracts and the Theory of the Firm. *The Quarterly Journal of Economics*, 117(1), 39-84.
- Baker, H. K. & Haslam, J. A. (1973). Informational Needs of Individual Investors. *The Journal of Accountancy*, November, 64-69.
- Bale, M. & Dale, T. (1998). Public Sector Reform in New Zealand and Its Relevance to Developing Countries. *The World Bank Research Observer*, 13(1), 103-121.
- Banker, R. D., Datar, S. M. & Maindiratta, A. (1989). Unobservable outcomes and multiattribute preferences in the evaluation of managerial performance. *Contemporary Accounting Research*, 5(1), 96-124.
- Banks, W., Fisher, J. & Nelson, M. (1997). University Accountability in England, Wales and Northern Ireland: 1992-1994. *Journal of International Accounting, Auditing & Taxation*, 6(2), 211-227.
- Banks, W. & Nelson, M. (1994). Financial Disclosures by Ontario Universities: 1998-1993. *Journal of International Accounting, Auditing & Taxation*, 3(2), 287-305.
- Beckett, R. & Gilmore, S. (2007). Lifting the Standard of Public Sector Performance Reporting. *Chartered Accountant Journal*, November 2007, 12-15.
- Behn, R. D. (1995). The Big Questions of Public Administration. *Public Administration Review*, 55(4), 313-324.
- Belkaoui, A. & Karpik, P. G. (1989). Determinants of the Corporate Decision to Disclose Social Information. *Accounting, Auditing & Accountability Journal*, 2(1), 36-51.



- Berelson, B. (1952). *Content Analysis in Communications Research*. Glencoe, Ill.: The Free Press.
- Bevir, M., Rhodes, R. A. W. & Weller, P. (2003). Traditions of Governance: Interpreting the Changing Role of the Public Sector. *Public Administration*, 81(1), 1-17.
- Boston, J. (1992). Assessing the Performance of Departmental Chief Executives: Perspectives from New Zealand. *Public Administration*, 70(3), 405-428.
- Boston, J., Martin, J., Pallot, J. & Walsh, P. (1996). *Public Management: The New Zealand Model*. Auckland: Oxford University Press.
- Boston, J. & Pallot, J. (1997). Linking Strategy and Performance: Developments in the New Zealand Public Sector. *Journal of Policy Analysis and Management*, 16(3), 382-404.
- Boyne, G. & Law, J. (1991). Accountability and Local Authority Annual Reports: The Case of Welsh District Councils. *Financial Accountability & Management*, 7(3), 179-194.
- Boyne, G. & Law, J. (2005). Setting Public Service Outcome Target: Lessons from Local Public Service Agreements. *Public Money & Management*, 25(4), 253-260.
- Bozzolan, S., Favotto, F. & Ricceri, F. (2003). Italian Annual Intellectual Capital Disclosure. *Journal of Intellectual Capital*, 4(4), 543-558.
- Brennan, N. (2001). Reporting Intellectual Capital in Annual Reports: Evidence from Ireland. *Accounting, Auditing & Accountability Journal*, 14(4), 423-436.
- Brun, M. E. & Siegel, J. P. (2006). What does appropriate performance reporting for political decision makers require? *International Journal of Productivity and Performance Management*, 55(6), 480-497.
- Brusca Alijarde, M. I. (1997). The Usefulness of Financial Reporting in Spanish Local Governments. *Financial Accountability & Management*, 26(3), 17-34.
- Bryan, S. H. (1997). Incremental Information Content of Required Disclosures Contained in Management Discussion and Analysis. *The Accounting Review*, 72(2), 285-301.
- Bukh, P. N., Nielsen, C., Gormsen, P. & Mouritsen, J. (2005). Disclosure of Information on Intellectual Capital in Danish IPO Prospectuses. *Accounting, Auditing & Accountability Journal*, 18(6), 713-732.
- Buzby, S. L. (1974). Selected Items of Information and their Disclosure in Annual Reports. *The Accounting Review*, 49(3), 423-435.
- Campbell, C. (2001). Juggling Inputs, Outputs, and Outcomes in the Search for Policy Competence: Recent Experience in Australia. *Governance: An International Journal of Policy and Administration*, 14(2), 253-282.

- Carlin, T. M. (2004). Output-Based Management and the Management of Performance: Insights from the Victorian Experience. *Management Accounting Research*, 15, 267-283.
- Carter, N. (1991). Learning to Measure Performance: The Use of Indicators in Organizations. *Public Administration*, 69(1), 85-101.
- Catusus, B. & Grondlund, A. (2005). More peace for less money: measurement and accountability in the Swedish Armed Forces. *Financial Accountability & Management*, 21(4), 467-484.
- Chapman, L. (1979). *Your Disobedient Servant*. Harmondsworth: Penguin.
- Charmichael, H. L. (1989). Self-Enforcing Contracts, Shirking, and Life Cycle Incentives. *Journal of Economic Perspectives*, 3(4), 65-83.
- Choi, F. D. S. (1973). Financial Disclosure and Entry to the European Capital Market. *Journal of Accounting Research*, 11(2), 533-541.
- Chow, C. W. & Wong-Boren, A. (1987). Voluntary Financial Disclosures by Mexican Corporations. *The Accounting Review*, 62(3), 533-541.
- Christensen, P. O. & Feltham, G. A. (2005). *Economics of Accounting: Volume II - Performance Evaluation*. New York, NY: Springer.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(Winter), 37-46.
- Collier, P. M. (2006). In Search of Purpose and Priorities: Police Performance Indicators in England and Wales. *Public Money & Management*, 26(3), 165-172.
- Controller and Auditor General. (1999). Third Report for 1999. Wellington.
- Controller and Auditor General. (2008). The Auditor-General's observations on the quality of performance reporting.
- Cook, A.-L. (2004). "Managing for Outcomes" in the New Zealand Public Management System, *New Zealand Treasury Working Paper 04/15*. Wellington: New Zealand Treasury.
- Cowan, S. & Gadenne, D. (2005). Australian Corporate Environmental Reporting: A Comparative Analysis of Disclosure Practices across Voluntary and Mandatory Disclosure Systems. *Journal of Accounting & Organizational Change*, 1(2), 165-179.
- Cowen, P. B., Cowen, T. & Tabarrok, A. (1992). An Analysis of Proposals for Constitutional Change in New Zealand. Wellington: New Zealand Business Roundtable.
- Cowen, S. S., Ferreri, L. B. & Parker, L. D. (1987). The Impact of Corporate Characteristics on Social Responsibility Disclosure: A Typology and Frequency-Based Analysis. *Accounting, Organizations and Society*, 12(2), 111-122.

- Coy, D., Fisher, M. & Gordon, T. (2001). Public Accountability: A New Paradigm for College and University Annual Reports. *Critical Perspectives on Accounting*, 12(1), 1-31.
- Cunningham, G. M. & Harris, J. E. (2005). Toward a Theory of Performance Reporting to Achieve Public Sector Accountability: A Field Study. *Public Budgeting & Finance*, 25(2), 15-42.
- Day, P. & Klein, R. (1987). *Accountabilities. Five public services*. London: Tavistock.
- Dees, J. (1992). Principals, Agents and Ethics. In N. E. Bowie & R. E. Freeman (Eds.), *Ethics and Agency Theory: An Introduction* (pp. 25-58). New York, NY: Oxford University Press.
- Di Francesco, M. (1999). Measuring Performance in Policy Advice Output: Australian Developments. *The International Journal of Public Sector Management*, 12(5), 420-431.
- Dixit, A. (1997). Power of incentives in private versus public organizations. *AEA Papers and Proceedings, May*, 378-382.
- Dixon, J., Kouzmin, A. & Korac-Kakabadse, N. (1998). Managerialism - Something Old, Something Borrowed, Little New: Economic Prescriptions versus Effective Organizational Change in Public Agencies. *International Journal of Public Sector Management*, 11(2/3), 164-187.
- Dunleavy, P., Margetts, H., Bastow, S. & Tinkler, J. (2006). New Public Management is Dead - Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, 16(3), 467-494.
- Economics and Strategy Group. (2003). Departmental Uptake of the Managing for Outcomes Initiative [Electronic Version], from <http://www.ssc.govt.nz/departamental-uptake-mfo>
- Eichbaum, C. & Shaw, R. (2007). Ministerial Advisers, Politicization and the Retreat from Westminster: The Case of New Zealand. *Public Administration*, 85(3), 609-640.
- Eisenhardt, K. M. (1989). Agency theory: an assessment and review. *Academy of Management Review*, 14(1), 57-74.
- Eisner, M. A., Worsham, J. & Ringquist, E. (1996). Crossing the Organizational Void: The Limits of Agency Theory in the Analysis of Regulatory Control. *Governance: An International Journal of Policy and Administration*, 9(4), 407-428.
- Fearing, F. (1954). Human Communication. Dept. of Psychology, University of California.
- Finance and Expenditure Committee. (1989). *Report on the Public Finance Bill 1989*. Wellington, New Zealand. (G. Printer o. Document Number)
- Frant, H. (1996). High-Powered and Low-Powered Incentives in the Public Sector. *Journal of Public Administration Research and Theory*, 6(3), 365-381.

- Frost, G. & Wilmshurst, T. (2000). The Adoption of Environment Related Management Accounting: An Analysis of Corporate Environment Sensitivity. *Accounting Forum*, 24(4).
- Gibbons, R. & Murphy, K. (1992). Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *The Journal of Political Economy*, 100(3), 468-505.
- Glynn, J. & Perkins, D. (1997). Control and Accountability in the NHS Market: A Practical Proposition or Logical Impossibility? *International Journal of Public Sector Management*, 10(1/2), 62-75.
- Government Accounting Standards Board (GASB). (1987). Concepts Statement No. 1: Objectives of Financial Reporting.
- Gray, D. (2005). *Social Development: Focusing on Outcomes Monitoring Results Developments in New Zealand*. Paper presented at the the Canadian Social Welfare Policy Conference. Retrieved 5 March 2008, from <http://www.ccsd.ca/cswp/2005/managingforoutcomes.pdf>
- Gray, R. & Haslam, J. (1990). External Reporting by UK Universities: An Exploratory Study of Accounting Change. *Financial Accountability & Management*, 6(1), 51-72.
- Gray, R., Kouhy, R. & Lavers, S. (1995). Constructing a Research Database of Social and Environmental Reporting by UK Companies. *Accounting, Auditing & Accountability Journal*, 8(2), 78-101.
- Gregory, R. (1998). A New Zealand Tragedy: Problems of Political Responsibility. *Governance: An International Journal of Policy and Administration*, 11(2), 231-240.
- Gregory, R. (2003). All the king's horses and all the king's men: Putting New Zealand's public sector back together again. *International Public Management Review*, 4(2), 41-58.
- Greve, C. (2000). Exploring contracts as reinvented institutions in the Danish public sector. *Public Administration*, 78(1), 153-164.
- Guthrie, J. & Abeysekera, I. (2006). Content Analysis of Social Environmental Reporting: What is New? *Journal of Human Resource Accounting & Accounting*, 10(2), 114-126.
- Guthrie, J. & Parker, L. D. (1990). Corporate social disclosure practice: A comparative international analysis. *Advances in Public Interest Accounting*, 3, 159-175.
- Guthrie, J., Petty, R., Yongvanich, K., & Ricceri, F. (2004). Using Content Analysis as a Research Method to Inquire into Intellectual Capital Reporting. *Journal of Intellectual Capital*, 5(2), 282-293.
- Hackston, D. & Milne, M. J. (1996). Some determinants of social and environmental disclosures in New Zealand companies. *Accounting, Auditing & Accountability Journal*, 9(1), 77-108.

- Hassan, O. & Marston, C. L. (2010). Disclosure Measurement in the Empirical Accounting Literature - A Review Article. Department of Economics and Finance, School of Sciences, Brunel University.
- Hendry, J. (2002). The principal's other problems: honest incompetence and the specification of objectives. *Academy of Management Review*, 27(1), 98-113.
- Herawaty, M. & Hoque, Z. (2007). Disclosure in the annual reports of Australian government departments. *Journal of Accounting & Organizational Change*, 3(2), 147-168.
- Ho, A. T.-K. & Ni, A. U. (2005). Have Cities Shifted to Outcome-Oriented Performance Reporting? - A Content Analysis of City Budgets. *Public Budgeting & Finance*, 25(2), 61-83.
- Ho, L.-C. J. & Taylor, M. E. (2007). An Empirical Analysis of Triple Bottom-Line Reporting and its Determinants: Evidence from the United States and Japan. *Journal of International Financial Management and Accounting*, 18(2), 123-150.
- Holsti, O. R. (1968). Content Analysis. In G. Lindzey & E. Aronson (Eds.), *The Handbook of Social Psychology* (Vol. 2). Reading, Mass.: Addison-Wesley.
- Hood, C. (1995). The "New Public Management" in the 1980s: Variations on a Theme. *Accounting, Organizations and Society*, 20(2/3), 93-109.
- Hooks, J., Coy, D. & Davey, H. (2001). The Annual Reports of New Zealand Electricity Companies: Assessing Quality. *Pacific Accounting Review*, 13(2), 35-69.
- Hoque, Z., Arends, S. & Alexander, R. (2004). Policing the Police Service: A Case Study of the Rise of 'New Public Management' within an Australian Police Service. *Accounting, Accountability & Performance*, 17, 59-84.
- Hyndman, N. & Anderson, R. (1995). The Use of Performance Information in External Reporting: An Empirical Study of UK Executives Agencies. *Financial Accountability & Management*, 11(1), 1-17.
- Hyndman, N. & Anderson, R. (1998). Performance Information, Accountability and Executive Agencies. *Public Money & Management*, 18(2), 23-30.
- Hyndman, N. & Eden, R. (2001). Rational Management, Performance Targets and Executive Agencies: Views from Agency Chief Executives in Northern Ireland. *Public Administration*, 79(3), 579-598.
- Hyndman, N. & Eden, R. (2002). Executive Agencies, Performance Targets and External Reporting. *Public Money & Management*, 22(3), 17-24.
- Ingram, R. W. & Frazier, K. B. (1980). Environmental Performance and Corporate Disclosure. *Journal of Accounting Research*, 18(2), 614-622.
- Jacobides, M. G. & Croson, D. C. (2001). Information policy: shaping the value of agency relationship. *Academy of Management Review*, 26(2), 202-223.

- James, O., & John, P. (2006). Public Management in the Ballot Box: Performance Information and Electoral Support for Incumbent English Local Governments. *Journal of Public Administration Research and Theory*, 17, 567-580.
- Johnson, R. W. M. (1994). The National Interest, Westminster, and Public Choice. *Australian Journal of Agricultural Economics*, 38(1), 1-30.
- Jones, D., Scott, R., Kimbro, L., & Ingram, R. W. (1985). *GASB Research Report 1: The Needs of Users of Governmental Financial Report*. Stamford: Governmental Accounting Standards Board.
- Jones, M. J. & Shoemaker, P. A. (1994). Accounting Narratives: A Review of Empirical Studies of Content and Readability. *Journal of Accounting Literature*, 1994(13).
- Julnes, P. d. L. & Holzer, M. (2001). Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation. *Public Administration Review*, 61(6), 693-708.
- Kassarjian, H. H. (1977). Content Analysis in Consumer Research. *The Journal of Consumer Research*, 4(1), 8-18.
- Kolbe, R. H. & Burnett, M. S. (1991). Content-Analysis Research: An Examination of Applications with Directives for Improving Research Reliability and Objectivity. *The Journal of Consumer Research*, 18(2), 243-250.
- Krippendorff, K. (1980). *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA: Sage Publishing Co.
- Laegreid, P., Roness, P. G. & Rubecksen, K. (2006). Performance management in practice: the Norwegian way. *Financial Accountability & Management*, 22(3251-270).
- Lambert, R. A. (2001). Contracting theory and accounting. *Journal of Accounting and Economics*, 32, 3-87.
- Lawton, A., McKeivitt, D. & Millar, M. (2000). Coping with Ambiguity: Reconciling External Legitimacy and Organizational Implementation in Performance Measurement. *Public Money & Management*, 20(3), 13-19.
- Lee, J. (2008). Preparing Performance Information in the Public Sector: An Australian Perspective. *Financial Accountability & Management*, 24(2), 117-149.
- Leventis, S. & Weetman, P. (2004). Impression Management: Dual Language Reporting and Voluntary Disclosure. *Accounting Forum*, 28, 307-328.
- Liu, X. & Mills, A. (2007). Agency theory and its applications in health care. In A. S. Preker, X. Liu, E. V. Velenyi & E. Baris (Eds.), *Public Ends, Private Means: Strategic Purchasing of Health Services* (pp. 151-178). Washington, D.C.: The World Bank.
- Mack, J. & Ryan, C. (2006). Reflections on the Theoretical Underpinnings of the General-Purpose Financial Reports of Australian Government Departments. *Accounting, Auditing & Accountability Journal*, 19(4), 592-612.

- Marston, C. L. & Shrides, P. J. (1991). The Use of Disclosure Indices in Accounting Research: A Review Article. *British Accounting Review*, 23, 195-210.
- Martimort, D. (1996). The organization of government: The multiprincipal nature of government. *European Economic Review*, 40, 673-685.
- Mayston, D. J. (1985). Non-Profit Performance Indicators in the Public Sector. *Financial Accountability & Management*, 1(1), 51-74.
- Mayston, D. J. (1993). Principals, Agents and the Economics of Accountability in the New Public Sector. *Accounting, Auditing & Accountability Journal*, 6(3), 68-96.
- McClelland, D. (1960). *The achieving society*. Princeton, NJ: Van Nostrand.
- McCubbins, M. D. & Schwartz, T. (1984). Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science*, 29, 721-748.
- Milgrom, P. & Roberts, J. (1992). *Economics, Organization, and Management*. New Jersey: Prentice Hall.
- Milne, M. J. & Adler, R. W. (1999). Exploring the Reliability of Social and Environmental Disclosures Content Analysis. *Accounting, Auditing & Accountability Journal*, 12(2), 237-256.
- Moe, T. M. (1995). The Politics of Structural Choice: Topward a Theory of Public Bureaucracy. In O. E. Williamson (Ed.), *Organization Theory: From Chester Barnard to the Present and Beyond* (Expanded ed.). New York: Oxford University Press.
- Moe, T. M. (1997). The Positive Theory of Public Bureaucracy. In D. C. Mueller (Ed.), *Perspective on Public Choice: A Handbook* (pp. 455-480). New York: Cambridge University Press.
- Mulgan, R. (2008). Public Sector Reform in New Zealand: Issues of Public Accountability. *Public Administration Quarterly*, 32(1), 1-31.
- Neale, A. & Anderson, B. (2000). *Performance reporting for accountability purposes - lessons, issues, future*. Paper presented at the International Public Management Workshop.
- Newberry, S. & Pallot, J. (2004). Freedom or Coercion? NPM Incentives in New Zealand Central Government Departments. *Management Accounting Research*, 15(3), 247-266.
- Niskanen, W. A. (1971). *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
- Norman, R. (2002). Managing through measurement or meaning? Lessons from experience with New Zealand's public sector performance management systems. *International Review of Administrative Sciences*, 68, 619-628.
- Ostrom, V. & Ostrom, E. (1971). Public Choice: A Different Approach to the Study of Public Administration. *Public Administration Review*, 31(2), 203-216.

- Parker, G. R. (2005). Reputational Capital, Opportunism, and Self-Policing in Legislature. *Public Choice*, 122(3/4), 333-354.
- Pendlebury, M., Jones, R. & Karbhari, Y. (1994). Developments in the Accountability and Financial Reporting Practices of Executive Agencies. *Financial Accountability & Management*, 10(1), 33-46.
- Petrie, M. & Webber, D. (2006). Review of evidence on broad outcome of public sector management regime. New Zealand Treasury.
- Poister, T. H. & Streib, G. (1999). Performance Measurement in Municipal Government: Assessing the State of the Practice. *Public Administration Review*, 59(4), 325-335.
- Pollanen, R. M. (2005). Performance Measurement in Municipalities: Empirical Evidence in Canadian Context. *International Journal of Public Sector Management*, 18(1), 4-24.
- Pollitt, C. (2006). Performance Information for Democracy: The Missing Link? *Evaluation*, 12(1), 38-55.
- Rantanen, H., Kulmala, H. I., Lonnqvist, A. & Kujansivu, P. (2007). Performance measurement systems in the Finnish public sector. *International Journal of Public Sector Management*, 20(5), 415-433.
- Ratanajongkol, S., Davey, H. & Low, M. (2006). Corporate Social Reporting in Thailand: The News is All Good and Increasing. *Qualitative Research in Accounting & Management*, 3(1), 67-83.
- Rhodes, R. A. W. (2005). Is Westminster Dead in Westminster (and Why Should We Care)? Australian National University.
- Robbins, W. A. & Austin, K. R. (1986). Disclosure Quality in Governmental Financial Reports: An Assessment of the Appropriateness of a Compound Measure. *Journal of Accounting Research*, 24(2), 412-421.
- Roberts, N. C. (2002). Keeping Public Officials Accountable through Dialogue: Resolving Paradox. *Public Administration Review*, 62(6), 658-669.
- Roberts, R. W. (1992). Determinants of Corporate Social Responsibility Disclosure: An Application of Stakeholder Theory. *Accounting, Organizations and Society*, 17(6), 595-612.
- Rutherford, B. A. (1996). The Structure and Content of Executive Agency Financial Statements. *Financial Accountability & Management*, 12(1), 1-20.
- Rutherford, B. A. (2000). The Construction and Presentation of Performance Indicators in Executive Agency External Reports. *Financial Accountability & Management*, 16(3), 225-249.
- Ryan, B. (2006). Beyond Westminster: Thinking the Aotearoa/New Zealand Way of Governing. *Policy Quarterly*, 2(3), 40-47.



- Ryan, C., Stanley, T. & Nelson, M. (2002). Accountability Disclosures by Queensland Local Government Councils: 1997-1999. *Financial Accountability & Management*, 18(3), 261-289.
- Schick, A. (1996). The spirit of reform: Managing the New Zealand state sector in a time of change. State Service Commission.
- Scott, G. (1996). *Government Reform in New Zealand*. Washington DC.
- Scott, G. (2001). *Public Management in New Zealand: Lessons and Challenges*. Wellington: New Zealand Business Roundtable.
- Scott, G., Ball, I. & Dale, T. (1997). New Zealand's Public Sector Management Reform: Implications for the United States. *Journal of Policy Analysis and Management*, 16(3), 357-381.
- Scott, G., Bushnell, P. & Sallee, N. (1990). Reform of the Core Public Sector: New Zealand Experience. *Governance: An International Journal of Policy and Administration*, 3(2), 138-167.
- Sharma, A. (1997). Professional as Agent: Knowledge Asymmetry in Agency Exchange. *The Academy of Management Review*, 22(3), 758-798.
- Singhvi, S. & Desai, H. B. (1971). An Empirical Analysis of the Quality of Corporate Financial Disclosure. *The Accounting Review*, 46(1), 129-138.
- Smith, M., & Taffler, R. (1995). The Incremental Effect of Narrative Accounting Information in Corporate Annual Reports. *Journal of Business Finance & Accounting*, 22(8), 1195-1210.
- Smith, M. & Taffler, R.J. (2000.) The Chairman's Statement - A Content Analysis of Discretionary Narrative Disclosures. *Accounting, Auditing & Accountability Journal*, 13(5), 624 - 647.
- Smith, P. (1995). On the Unintended Consequences of Publishing Performance Data in the Public Sector. *International Journal of Public Administration*, 18, 277-310.
- Smith, S. A. & Coy, D. (2000). The Quality of City Council Annual Reports, 1996-97 & 1997-98: Preliminary Findings, *Working Paper Series*: The University of Waikato.
- Somers, J.-C. (2008). *Bureaucracy, Democracy and Westminster: The Quest for Political Control*. Victoria University of Wellington, Wellington.
- Sotirakou, T. & Zeppou, M. (2006). Utilizing performance measurement to modernize the Greek public sector. *Management Decision*, 44(9), 1277-1304.
- State Services Commission. (1998a). A Better Focus on Outcomes through SRA Networks. Wellington: State Service Commission.
- State Services Commission. (1998b). *A Better Focus on Outcomes through SRA Networks*: State Services Commission. Document Number)

- State Services Commission. (2004). Public Service Induction - What it means to be a public servant [Electronic Version]. Retrieved 16 April 2010, from [http://www.ssc.govt.nz/display/document.asp?docid=6853&pageno=2#P27\\_2886](http://www.ssc.govt.nz/display/document.asp?docid=6853&pageno=2#P27_2886)
- Steccolini, I. (2004). Is the Annual Report an Accountability Medium? An Empirical Investigation into Italian Local Governments. *Financial Accountability & Management*, 20(3), 327-350.
- Steering Group. (1991). Review of state sector reforms. Wellington: State Services Commission.
- Steering Group for the Managing for Outcomes Roll-out 2004/05. (2003). Managing for Outcomes: Guidance for Departments. Wellington: Department of the Prime Minister and Cabinet, Te Puni Kokiri, State Services Commissions, & The Treasury.
- Stevens, D. E. (2002). The Effects of Reputation and Ethics on Budgetary Slack. *Journal of Management Accounting Research*, 14, 153-170.
- Stewart, J. & Walsh, K. (1994). Performance Measurement: When Performance can Never be Finally Defined. *Public Money & Management*, 14(2), 45-49.
- Stigler, G. J. (1971). The Theory of Economic Regulation. *The Bell Journal of Economics and Management Science*, 2(1), 3-21.
- Sydserrf, R. & Weetman, P. (2002). Developments in Content Analysis: A Transitivity Index and DICTION Scores. 2002, 15(4), 523-545.
- Taylor, D. & Rosair, M. (2000). The Effect of Participating Parties, the Public and Size on Government Departments' Accountability Disclosure in Annual Reports. *Accounting, Accountability & Performance*, 6(1), 77-97.
- The Pathfinder Project. (2003a). Building Block 2: Outcome Indicators [Electronic Version]. *Guidance on Outcomes Focused Management*. Retrieved 5 April 2008, from <http://io.ssc.govt.nz/pathfinder>
- The Pathfinder Project. (2003b). Building Block 3: Intervention Logic [Electronic Version]. *Guidance on Outcomes Focused Management*. Retrieved 5 April 2008, from <http://io.ssc.govt.nz/pathfinder>
- The Steering Group for the Managing for Outcomes Roll-out 2004/05. (2003). Managing for Outcomes: Guidance for Departments: Department of the Prime Minister and Cabinet, Te Puni Kokiri, State Services Commission, and the Treasury.
- The Treasury. (2006). Preparing the Annual Report. from <http://www.treasury.govt.nz/publicsector/annualreport/>
- The Treasury and State Services Commission. (2007). Guidance and Requirements for Departments: Preparing the Statement of Intent.

- Thomas, P. (1998). The Changing Nature of Accountability. In B. Peters & D. Savoie (Eds.), *Taking Stock: Assessing Public Sector Reforms*. Montreal: McGill-Queens University Press.
- Thompson, G. D. (1995). Problems with Service Performance Reporting: the Case of Public Art Galleries. *Financial Accountability & Management*, 11(4), 337-350.
- Tooley, S. & Guthrie, J. (2007). Reporting Performance by New Zealand Secondary School: An Analysis of Disclosures. *Financial Accountability & Management*, 23(4), 351-374.
- Tullock, G. & Eller, K. (1994). *Rent Seeking*. London: Edward Elgar.
- Ussher, C. & Kibblewhite, A. (2005). Outcome Focussed Management in New Zealand: A Background Paper. Wellington: The Treasury.
- Vandemaele, S. N., Vergauwen, P. G. M. C. & Smits, A. J. (2005). Intellectual Capital Disclosure in The Netherlands, Sweden and the UK. *Journal of Intellectual Capital*, 6(3), 417-426.
- Verbeeten, F. H. M. (2008). Performance management practices in public sector organizations: impact on performance. *Accounting, Auditing, & Accountability Journal*, 21(3), 427-454.
- Waterman, R. W. & Meier, K. J. (1998). Principal-agent models: an expansion. *Journal of Public Administration Research and Theory*, 8(2), 173-202.
- Waweru, N., Porporato, M. & Hoque, Z. (2007). Performance Measurement Practices in Canadian Government Departments: A Survey. Unpublished Paper. York University & La Trobe University.
- Webber, D. (2004). Wrestling with Outcomes: The New Zealand Experience. *Agenda*, 11(4), 335-348.
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.
- Wilson, J. Q. (1989). *Bureaucracy: What Government Agencies Do and Why They Do*. New York: Basic Books.
- Wright, P., Mukherji, A. & Kroll, M. J. (2001). A reexamination of agency theory assumptions: extension and extrapolations. *Journal of Socio-Economics*, 30, 413-429.
- Zarzeski, M. T. (1996). Spontaneous Harmonization Effects of Culture and Market Forces on Accounting Disclosure Practices. *Accounting Horizons*, 10(1), 18-37.

## APPENDIX A

### Selected Research on Public Sector Performance Information

Study by	Document analysed	Item of analysis	Methodology	Findings
Boyne and Law (1991)	Welsh District (UK) Council's SSP 1981/1982 – 1988/1999	Performance information in 9 categories	Simple scoring method	Poor quality, some measures of output and efficiency
Boyne & Law (2005)	Local authorities (UK)	Performance indicators	Categorization of performance indicator into 12 dimensions	Modest use of outcome indicators (49%)
Ho & Ni (2005)	21 cities in the US	Performance indicators	Indicator counting and categorisation	Significant progress in performance measurement and reporting. Few services remain primarily output-focused
Poister & Streib (1999)	US cities	Performance indicators	Indicator counting and survey	Substantial use of performance measures with some area exceptions. More use of workload and effectiveness measures, less use of unit cost and efficiency measures
Robbins & Austin (1986)	US cities	Annual reports	Disclosure quality index	Coalition of voters, administrative power and management incentives are influencing the quality of disclosure. No material impact on the use of compound disclosure index
Ryan et. al. (2002)	Queensland local government councils	Annual reports	Local Government Accountability index	The quality of reporting increased over time. Financial disclosure is more complete.

Study by	Document analysed	Item of analysis	Methodology	Findings
Smith & Coy (2000)	NZ local authorities	Annual reports	Accountability index	The quality of reporting only fulfils the statutory requirements
Thompson (1995)	NZ Public art galleries	Performance indicators	Indicator counting and categorisation	The essential requirements of a service performance report are not achieved
Tooley & Guthrie (2007)	NZ secondary schools	Annual reports	School Performance Accountability Disclosure Index	Many schools presented limited performance information.
Collier (2006)	Police agencies in England and Wales	Performance indicators	Indicator counting and categorisation	Changes in indicators from one initiative to another has not been incremental, unidirectional, and unambiguous.
Carter (1991)	Government departments, public agencies and private businesses	Performance indicator	Case studies: information characteristic related to organisational dimension	Cross-cutting organisational characteristics are most helpful in explaining variation in the nature and use of PIs.
Hyndman & Anderson (1995)	UK Executive Agencies	Performance indicators	Indicator counting and categorisation	Significant proportions of agencies reporting little or no information of efficiency and effectiveness
Hyndman & Eden (2002)	UK executive agencies	Performance targets and indicators	Target and measures counting	There is a high degree of quantification and coordination in the performance measurement and performance reporting systems
Hyndman & Anderson (1998)	UK executive agencies	Performance indicators	Indicator counting and categorisation of measures into several characteristics	Significant proportions of agencies reporting little or no information of efficiency and effectiveness

Study by	Document analysed	Item of analysis	Methodology	Findings
Pendlebury et al. (1994)	UK executive agencies	Performance indicators	Indicator counting and categorisation	Less accountability to Parliament and the public more accountability to the customers of goods and services
Rutherford (2000)	UK executive agencies	Performance indicators	Data counting based on several information characteristics	More emphasis on measures of efficiency and effectiveness A variety of different methods of presentation

## APPENDIX B

### Selected Accounting Research using a Disclosure Index

No.	Object of Analysis	Author(s)	Focus	Disclosure index or rating
1.	Annual reports	Brennan (2001)	The association between the difference of book to market value and the extent of intellectual capital asset disclosure	A framework developed by Guthrie et. al. (1999) for intellectual disclosure framework involves 24 variables across three intellectual capital categories.
2.	Annual reports (public sector)	Robbins & Austin (1986)	The determinants of disclosure quality in municipal annual reports	An index is developed by sending questionnaire to municipal bond analysts resulting 27 information items in the order of importance. Weighted scores are used.
3.	Annual reports	Belkaoui & Karpik (1989)	The relationship between the extent of social disclosure and some characteristics (social performance, economic performance, political visibility and contracting and monitoring costs)	An index derived from Ernst and Ernst surveys of social responsibility disclosure (Ernst and Ernst, 1973) transformed into a scale, varies from 0 to 13, to represent the number of social responsibility programs.
4.	Annual reports	Cowen et al. (1987)	The relationship between a number of corporate characteristics and specific types of social responsibility disclosures	An index derived from Ernst and Whinney 1978 survey of corporate social responsibility disclosure. The index has seven categories: environment, energy, fair business practices, human resources, community involvement, product safety and other disclosures).
5.	Annual reports, stand-alone reports and special website reports	Ho & Taylor (2007)	The determinants of triple bottom-line disclosures.	Sixty disclosure items are selected to determine the extent of economic, social, and environmental reporting. The sources: 2002 GRI Reporting Guidelines, Deloitte and Touche (2002), KPMG (2002) and Pacific Sustainability Index (PSI) Scoring Sheet and some literatures.
6.	IPO prospectuses	Bukh et al. (2005)	The relationship between the extent of intellectual capital disclosure and the firm's characteristics	A disclosure index is developed based on a thorough inspection of the literature on corporate disclosure and intellectual capital reporting. There are 78 items in the index.

No.	Object of Analysis	Author(s)	Focus	Disclosure index or rating
7.	Annual reports	Leventis & Weetman (2004)	The impact of dual reporting on the extent of voluntary disclosures	A checklist based on the structure by Gray et. al. (1995) and Meek et. al. (1995). Some items are added based on theoretical and empirical supports. The list contains three disclosure categories: corporate environment (26 items), social responsibility (19 items) and financial (27 items).
8.	Management Discussion and Analysis (MDAs)	Bryan (1997)	The association between MD&A variables and future financial variables, the association between MD&A and Value-Line Sales Forecast Revision, and the association between MD&A and stock returns.	An index is developed based on the disclosures required for the MD&A. Each disclosure is classified by its direction (unfavourable, neutral, favourable or missing)



## **APPENDIX C: List of Items in the Disclosure Index**

### **Output Information**

1. Output Description
  - a. Description of output class
  - b. A list of detailed outputs
  - c. Description of detailed outputs
  - d. Explanation of output class change
2. Cost disclosure
  - a. Total actual cost for each output class
  - b. Total planned cost for each output class
  - c. Cost allocation method disclosure
  - d. Variance disclosure
3. Presentation of quantity measures
  - a. Actual quantity (list of actual deliverables for binary-type output)
  - b. Planned quantity (list of planned deliverables for binary-type output)
  - c. Variance disclosure
4. Quality description
  - a. List of relevant quality dimensions
  - b. Descriptions of relevant quality dimensions
5. Quality benchmark used
  - a. Subjective (internal) benchmark
  - b. External benchmark
  - c. External assessment of external benchmark
6. Presentation of quality measures
  - a. Actual quality measures
  - b. Planned quality measures
  - c. Variance disclosure
7. Presentation of timeliness measures
  - a. Actual timeliness measure
  - b. Planned timeliness measure
  - c. Variance disclosure
8. Presentation of cost measures
  - a. Actual cost measure
  - b. Planned cost measure
  - c. Variance disclosure
9. Presentation of location measures
  - a. Actual location measure
  - b. Planned location measure
  - c. Variance disclosure

## Outcome Information

1. Goals/objectives identification
  - a. Description of goals/objectives
  - b. Link between goals/objectives and outcomes in graphical/tabular form
  - c. Elaborate explanation of the links between goals/objectives and outcomes
2. Outcome description
  - a. Title or short description of outcomes
  - b. Explanation of each outcomes
  - c. Outcome framework outlining high-level and low-level outcomes
  - d. Explanation of the links between high-level and low-level outcomes
3. Outcome target
  - a. Broad statement of targets
  - b. Quantitative targets
  - c. Multi-dimension targets
4. Outcome-output link discussion
  - a. Simple linking statement
  - b. Explanations to justify the links (intervention logic)
  - c. Discussion of output alternative
5. Discussion of other factors influencing outcomes
  - a. Identification of the factors
  - b. Explanation of the factors
  - c. Management assessment of projections of the factors
6. Presentation of outcome measures
  - a. Planned outcomes
  - b. Variance discussion
  - c. Past outcomes (measures from previous periods)
  - d. Current outcomes (measures from current periods)
  - e. Comparative outcome data from other organisations and/or countries
7. Cost effectiveness of the interventions
  - a. Descriptions of cost effectiveness
  - b. Explanation of measures of cost effectiveness
  - c. Actual measures
  - d. Planned measures
8. Organisation health and capability to deliver outcomes
  - a. Descriptions of organisation health and capability
  - b. Explanation of measures of organisation health and capability
  - c. Actual measures
  - d. Planned measures

## **APPENDIX D:**

### **List of departments selected for disclosure analysis**

1. Archives New Zealand (Te Whare Tohu Tuhituhinga O Aotearoa)
2. Crown Law Office
3. Department of Building and Housing (Te Tari Kaupapa Whare)
4. Department of Corrections (Ara Poutama Aotearoa)
5. Department of Internal Affairs (Te Tari Taiwhenua)
6. Department of Labour (Te Tari Mahi)
7. Inland Revenue Department (Te Tari Taake)
8. Land Information New Zealand (Toitu te whenua)
9. Ministry of Agriculture and Forestry (Te Manatū Ahuwhenua Ngāherehere)
10. Ministry of Defence (Manatu Kaupapa Waonga)
11. Ministry of Economic Development (Manatū Ōhanga)
12. Ministry for the Environment (Manatū Mō Te Taiao)
13. Ministry of Fisheries (Te Tautiaki i nga tini a Tangaroa)
14. Ministry of Foreign Affairs and Trade (Te Manatū Aorere)
15. Ministry of Health (Manatū Hauora)
16. Ministry of Justice (Tāhū o te Ture)
17. Ministry of Research, Science and Technology (MoRST) (Te Manatū Pūtaiao)
18. Ministry of Social Development (Te Manatu Whakahiato Ora)
19. Ministry of Women's Affairs (Minitatanga mō ngā Wāhine)
20. New Zealand Customs Service (Te Mana Arai o Aotearoa)
21. State Services Commission (Te Komihana O Ngā Tari Kāwanatanga)
22. Statistics New Zealand (Te Tari Tatau)
23. Te Puni Kōkiri (Ministry of Māori Development)
24. The Treasury (Kaitohutohu Kaupapa Rawa)
25. New Zealand Defence Force (NZDF) (Te Ope Kaatua O Aotearoa)
26. Parliamentary Counsel Office (New Zealand) (Te Tari Tohutohu Paremata)
27. New Zealand Police (Ngā Pirihimana O Aotearoa)

### **List of departments that are excluded due to unavailability of accountability documents for the period of observation**

1. Department of Conservation (Te Papa Atawhai)
2. Education Review Office (Te Tari Arotake Matauranga)
3. Government Communications Security Bureau
4. Department of the Prime Minister and Cabinet
5. Ministry of Education (Te Tāhuhu o te Mātauranga)
6. Ministry of Pacific Island Affairs
7. Ministry of Transport (Te Manatū Waka)
8. National Library of New Zealand (Te Puna Mātauranga o Aotearoa)
9. Serious Fraud Office (Te Tari Hara Taware)
10. New Zealand Security Intelligence Service
11. Office of the Clerk of the House of Representatives
12. Parliamentary Service Commission
13. New Zealand Food Safety Authority (Te Pou Oranga Kai O Aotearoa) (New Department)
14. Ministry for Culture and Heritage (Te Manatū Taonga) (New Department)

## APPENDIX E

**TABLE 1 Relative scores for output information**

	AGENCY	OUTPUT INFORMATION SCORE				
		2003/2004	2004/2005	2005/2006	2006/2007	2007/2008
1	New Zealand Defence Force	0.81	0.83	0.82	0.83	0.87
2	Ministry of Economic Development	0.78	0.84	0.83	0.82	0.81
3	Ministry of Health	0.79	0.80	0.81	0.81	0.78
4	Ministry of Research, Science and Technology	0.79	0.78	0.80	0.78	0.80
5	The Treasury	0.80	0.80	0.80	0.80	0.76
6	Department of Internal Affairs	0.75	0.79	0.80	0.78	0.77
7	Ministry of Defence	0.76	0.77	0.79	0.79	0.74
8	Land Information New Zealand	0.79	0.77	0.73	0.76	0.72
9	Inland Revenue Department	0.77	0.77	0.72	0.80	0.71
10	Ministry of Fisheries	0.74	0.75	0.76	0.76	0.73
11	Ministry of Social Development	0.79	0.71	0.71	0.74	0.72
12	Ministry of Foreign Affairs and Trade	0.71	0.72	0.72	0.76	0.74
13	New Zealand Customs Service	0.74	0.74	0.69	0.74	0.68
14	Parliamentary Counsel Office	0.68	0.67	0.67	0.67	0.85
15	Crown Law Office	0.73	0.72	0.70	0.69	0.67
16	Ministry of Justice	0.69	0.65	0.69	0.72	0.68
17	Department of Correction	0.70	0.70	0.68	0.69	0.62
18	Statistics New Zealand	0.65	0.63	0.67	0.68	0.68
19	State Services Commission	0.61	0.63	0.64	0.66	0.71
20	New Zealand Police	0.64	0.69	0.64	0.65	0.61
21	Archive New Zealand	0.66	0.60	0.65	0.62	0.62
22	Te Puni Kokiri	0.48	0.66	0.69	0.67	0.51
23	Ministry of Housing	0.58	0.48	0.50	0.71	0.70
24	Ministry of Women's Affairs	0.62	0.59	0.59	0.59	0.54
25	Department of Labour	0.53	0.58	0.54	0.51	0.50
26	Ministry of Agriculture and Forestry	0.63	0.38	0.45	0.42	0.44
27	Ministry for the Environment	0.42	0.31	0.27	0.34	0.38
Mean		0.69	0.68	0.68	0.70	0.68
Median		0.71	0.71	0.69	0.72	0.71
Maximum		0.81	0.84	0.83	0.83	0.87
Minimum		0.42	0.31	0.27	0.34	0.38
Standard deviation		0.10	0.13	0.13	0.12	0.12

**TABLE 2 Average score for each disclosure item – output information**

DISCLOSURE ITEMS		03/04	04/05	05/06	06/07	07/08
<b>A Output Description</b>						
1	Description of output class	1.00	1.00	1.00	1.00	1.00
2	A list of detailed outputs	0.96	0.93	0.93	0.93	0.93
3	Description of detailed outputs	0.96	0.93	0.89	0.93	0.93
4	Explanation of output class change	0.00	0.60	0.40	0.25	0.38
<b>B Cost disclosure</b>						
5	Total actual cost for each output class	0.96	0.96	0.99	0.99	0.96
6	Total planned cost for each output class	0.99	0.99	0.99	0.99	0.98
7	Cost allocation method disclosure	0.00	0.04	0.00	0.00	0.00
8	Variance disclosure	0.19	0.37	0.34	0.42	0.35
<b>C Presentation of quantity measures</b>						
9	Actual quantity (list of actual deliverables for binary-type output)	0.93	0.91	0.92	0.91	0.90
10	Planned quantity (list of planned deliverables for binary-type output)	0.73	0.71	0.69	0.70	0.71
11	Variance disclosure	0.53	0.48	0.52	0.54	0.54
<b>D Quality description</b>						
12	List of relevant quality dimensions	0.90	0.87	0.87	0.92	0.92
13	Descriptions of relevant quality dimensions	0.74	0.71	0.71	0.73	0.75
<b>E Quality benchmark used</b>						
14	Subjective (internal) benchmark	0.74	0.73	0.74	0.75	0.77
15	External benchmark	0.61	0.59	0.61	0.63	0.61
16	External assessment of external benchmark	0.60	0.58	0.60	0.62	0.60
<b>F Presentation of quality measures</b>						
17	Actual quality measures	0.77	0.78	0.81	0.86	0.86
18	Planned quality measures	0.78	0.76	0.77	0.82	0.78
19	Variance disclosure	0.87	0.81	0.83	0.84	0.68
<b>G Presentation of timeliness measures</b>						
20	Actual timeliness measure	0.73	0.69	0.70	0.70	0.71
21	Planned timeliness measure	0.65	0.64	0.63	0.63	0.64
22	Variance disclosure	0.78	0.63	0.67	0.71	0.76
<b>H Presentation of cost measures</b>						
23	Actual cost measure	0.16	0.17	0.13	0.17	0.13
24	Planned cost measure	0.21	0.19	0.15	0.19	0.11
25	Variance disclosure	0.10	0.05	0.03	0.21	0.20
OUTPUT INFORMATION - TOTAL SCORE		14.88	14.71	14.60	15.03	14.89

**TABLE 3 Relative scores for outcome information**

	AGENCY	OUTCOME INFORMATION SCORE				
		2003/2004	2004/2005	2005/2006	2006/2007	2007/2008
1	Department of Correction	0.41	0.45	0.41	0.55	0.51
2	Ministry of Health	0.34	0.52	0.50	0.45	0.41
3	Inland Revenue Department	0.45	0.52	0.33	0.47	0.47
4	The Treasury	0.38	0.45	0.45	0.42	0.47
5	Ministry of Housing	0.38	0.38	0.38	0.45	0.41
6	Archive New Zealand	0.31	0.31	0.30	0.39	0.62
7	New Zealand Customs Service	0.34	0.34	0.28	0.34	0.38
8	Department of Labour	0.38	0.38	0.31	0.31	0.31
9	Department of Internal Affairs	0.28	0.28	0.34	0.38	0.38
10	Ministry of Justice	0.31	0.31	0.31	0.34	0.38
11	Ministry of Foreign Affairs and Trade	0.21	0.24	0.34	0.41	0.41
12	Ministry of Agriculture and Forestry	0.24	0.31	0.31	0.31	0.44
13	New Zealand Defence Force	0.21	0.28	0.34	0.38	0.38
14	Statistics New Zealand	0.17	0.28	0.34	0.34	0.45
15	Ministry of Social Development	0.17	0.35	0.35	0.35	0.35
16	New Zealand Police	0.30	0.29	0.30	0.32	0.33
17	Parliamentary Counsel Office	0.24	0.24	0.28	0.31	0.31
18	Ministry of Research, Science and Technology	0.17	0.43	0.43	0.17	0.17
19	Ministry of Economic Development	0.24	0.24	0.28	0.28	0.28
20	Ministry of Fisheries	0.17	0.21	0.24	0.34	0.34
21	State Services Commission	0.34	0.34	0.17	0.17	0.21
22	Land Information New Zealand	0.21	0.21	0.24	0.28	0.28
23	Te Puni Kokiri	0.07	0.14	0.14	0.17	0.28
24	Ministry for the Environment	0.14	0.14	0.14	0.14	0.18
25	Ministry of Defence	0.10	0.14	0.14	0.14	0.17
26	Ministry of Women's Affairs	0.00	0.07	0.17	0.24	0.14
27	Crown Law Office	0.14	0.10	0.07	0.14	0.14
Mean		0.25	0.29	0.29	0.32	0.34
Median		0.24	0.29	0.31	0.34	0.35
Maximum		0.45	0.52	0.50	0.55	0.62
Minimum		0.00	0.07	0.07	0.14	0.14
Standard deviation		0.11	0.12	0.10	0.11	0.12

**TABLE 4 Average score for each disclosure item – output information**

DISCLOSURE ITEMS	03/04	04/05	05/06	06/07	07/08
<b>A Goals/objectives identification</b>					
1 Description of goals/objectives	0.78	0.81	0.70	0.74	0.74
2 Link between goals/objectives and outcomes in graphical/tabular form	0.59	0.56	0.59	0.63	0.67
3 Elaborate explanation of the links between goals/objectives and outcomes	0.00	0.00	0.00	0.00	0.00
<b>B Outcome description</b>					
4 Title or short description of outcomes	0.96	1.00	1.00	1.00	1.00
5 Explanation of each outcomes	0.74	0.85	0.89	0.96	0.96
6 Outcome framework outlining high-level and low-level outcomes	0.59	0.78	0.81	0.81	0.78
7 Explanation of the links between high-level and low-level outcomes	0.07	0.11	0.04	0.00	0.00
<b>C Outcome target</b>					
8 Broad statement of targets	0.48	0.56	0.70	0.74	0.82
9 Quantitative targets	0.09	0.21	0.14	0.14	0.20
10 Multi-dimension targets	0.36	0.54	0.54	0.59	0.74
<b>D Outcome-output link discussion</b>					
11 Simple linking statement (tabular or graphical form)	0.88	0.89	0.85	0.89	0.89
12 Explanations to justify the links (intervention logic)	0.04	0.15	0.19	0.15	0.15
13 Discussion of output alternative	0.00	0.00	0.00	0.00	0.00
<b>E Discussion of other factors influencing outcomes</b>					
14 Identification of the factors	0.26	0.26	0.33	0.34	0.37
15 Explanation of the factors	0.07	0.07	0.07	0.15	0.07
16 Management assessment of projections of the factors	0.00	0.00	0.00	0.00	0.00
<b>F Presentation of outcome measures</b>					
17 Planned outcomes	0.00	0.00	0.00	0.04	0.04
18 Variance discussion	n/a	n/a	n/a	n/a	n/a
19 Past outcomes (measures from previous periods)	0.14	0.21	0.20	0.16	0.17
20 Current outcomes (measures from current periods)	0.18	0.24	0.18	0.12	0.15
21 Comparative outcome data from other organisations and/or countries	0.00	0.03	0.03	0.00	0.00
<b>G Cost effectiveness of the interventions</b>					
22 Descriptions of cost effectiveness	0.04	0.04	0.07	0.26	0.42
23 Explanation of measures of cost effectiveness	0.04	0.04	0.00	0.11	0.16
24 Actual measures	0.00	0.00	0.00	0.07	0.11
25 Planned measures	0.00	0.00	0.00	0.00	0.00
<b>H Organisation health and capability to deliver outcomes</b>					
26 Descriptions of organisation health and capability	0.67	0.78	0.81	0.93	0.89
27 Explanation of measures of organisation health and capability	0.04	0.00	0.00	0.22	0.22
28 Actual measures	0.00	0.00	0.00	0.00	0.04
29 Planned measures	0.00	0.00	0.00	0.00	0.00
OUTCOME INFORMATION - TOTAL SCORE	7.24	8.38	8.35	9.24	9.80

## APPENDIX F

### Disclosure Index Research Instrument - Scoring Instruction

#### SCORING INSTRUCTION FOR INFORMATION AT OUTPUT CLASS LEVEL

The scoring is applied to the Statement of Intent and the corresponding Annual Report. The coder should check both documents in following the instructions below.

Items	Instruction	Control
Description of output classes	Identify the output classes and list them horizontally in the scoring sheet for output class level. This information is usually presented in the Statement of Intent and/or Annual Report.	
	For each output class identify, put a score of 1 in the appropriate cell if there is a description of output class. This description should be more than the name or title of output class. This information is usually presented in the Statement of Intent and/or Annual Report.	
A list of detailed outputs	For each output class, put a score of 1 in the appropriate cell if there is a list of names/titles for detailed outputs. The names/titles may be presented in a table. This information is usually presented in the Statement of Intent and/or Annual Report.	
Description of detailed outputs	For each output class, put a score of 1 in the appropriate cell if there are descriptions of detailed outputs. These descriptions should be more than the name/title of output. This information is usually presented in the Statement of Intent and/or Annual Report.	
Explanation of output class changes	Compare this year's SOI with last year's SOI. Check if there is/are an output class change(s). If an output class change is detected, identify if there is an explanation for the change. If yes, put a score of one in the cell that corresponding with the column of the changed output class. The change can be a modification or an addition to the output class. This information is usually presented in the Statement of Intent and/or Annual Report.	
Total actual cost for each output class	For each output class, put a score of 1 in the appropriate cell if the information on actual cost is presented. These descriptions should be more than the name/title of output. This information is usually presented in the Annual Report.	
Total planned cost for each output class	For each output class, put a score of 1 in the appropriate cell if the information on planned/budgeted cost is presented. These descriptions should be more than the name/title of output. This information is usually presented in the Statement of Intent and/or Annual Report.	



<b>Items</b>	<b>Instruction</b>	<b>Control</b>
Cost allocation method disclosure	Put a score of 1 in the appropriate cell on average column if there is a disclosure on cost allocation method. This information is usually presented in the Annual Report.	
Variance disclosure	For each output class, calculate if the variance between actual cost and planned/budgeted cost is 10% or more of the planned/budgeted cost. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for the variance. If there is an explanation, put a score of one in the appropriate cell. This information is usually presented in the Annual Report.	

## SCORING INSTRUCTION FOR INFORMATION AT OUTPUT LEVEL

The scoring is applied to the Statement of Intent and the corresponding Annual Report. The coder should check both documents in following the instructions below.

Items	Instructions	Control
	Identify the output and list them horizontally in the scoring sheet for output level. This information is usually presented in the Statement of Intent and/or Annual Report.	
Actual quantity for each output	For each output, put a score of 1 in the appropriate cell if the information on actual quantity is presented. This information is usually presented in the Annual Report. <ul style="list-style-type: none"> <li>• For binary-type output (output with a quantity of one), the actual quantity is deemed to be presented if there are a list of milestones/targets achieved or activities conducted.</li> </ul>	
Planned quantity for each output	For each output, put a score of 1 in the appropriate cell if the information on planned/budgeted quantity is presented. This information is usually presented in the Statement of Intent and/or Annual Report. <ul style="list-style-type: none"> <li>• For binary-type output (output with a quantity of one), the actual quantity is deemed to be presented if there are a list of planned milestones/targets or activities.</li> </ul>	
Quantity variance disclosure for each output	For each output, calculate if the variance between actual quantity and planned/budgeted quantity is 10% or more of the planned/budgeted quantity. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for the variance. If there is an explanation, put a score of 1 in the appropriate cell. This information is usually presented in the Annual Report. <ul style="list-style-type: none"> <li>• For binary-type output (output with a quantity of one), the variance is determined if the planned milestones/targets are not achieved or planned activities are not completed.</li> </ul>	
List of relevant quality dimensions	For each output, identify if there is a list of quality dimension is disclosed. If yes, put a score of 1 in the appropriate cell. <ul style="list-style-type: none"> <li>• The quality dimensions for policy advice outputs may be disclosed separately as a generic set of quality dimension. If is the case, then all policy advice outputs should get a score of 1.</li> </ul>	

Items	Instructions	Control
Description of quality dimensions	<p>For each output, identify if there is a description on the list of quality dimension is disclosed. If yes, put a score of 1 in the appropriate cell. A description means a presentation of quality dimension of more than the name/title of the quality dimension.</p> <ul style="list-style-type: none"> <li>The description of quality dimensions for policy advice outputs may be disclosed separately as a generic set of quality dimension. If is the case, then all policy advice outputs should get a score of 1.</li> </ul>	
Subjective (internal) benchmark	<p>For each output, identify in the list of quality dimension that is applicable to this output whether internal benchmarks are used. If yes, put a score of 1 in the appropriate cell. Internal benchmark is benchmark established by the entity being assessed.</p>	
External benchmark	<p>For each output, identify in the list of quality dimension that is applicable to this output whether internal benchmarks are used. If yes, put a score of 1 in the appropriate cell. Internal benchmark is benchmark established by entities/persons external to the organisation, including ministers, customers, etc.</p>	
External assessment of benchmark	<p>For each output, identify whether the quality dimensions are assessed by entities/persons external to organisation, including ministers, customers, etc. If yes, put a score of 1 in the appropriate cell. In the case there is no information regarding who made the assessment, it is assumed that the dimensions are assessed internally, so no score is given.</p>	
Actual quality measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on actual quality is presented. This information is usually presented in the Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the actual quality is deemed to be presented if at least there is a general statement on the result of quality assessment applicable for this type of outputs.</li> </ul>	
Planned quality measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on planned/budgeted quality is presented. This information is usually presented in the Statement of Intent and/or Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the planned quality is deemed to be presented if at least there is a general statement the expectation of the quality assessment for this type of outputs.</li> </ul>	
Quality variance disclosure for each output	<p>For each output, calculate if the variance between actual quality measures and planned/budgeted quality measures is 10% or more of the planned/budgeted quality measures. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for</p>	

Items	Instructions	Control
	<p>the variance. If there is an explanation, put a score of 1 in the appropriate cell. This information is usually presented in the Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the variance is determined if the expectation of the quality assessment differs from the result of the assessment.</li> </ul>	
Actual timeliness measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on timeliness measure is presented. This information is usually presented in the Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the actual timeliness measure is deemed to be presented if at least there is information about dates or deadlines of activities in the output.</li> </ul>	
Planned timeliness measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on planned/budgeted timeliness measure is presented. This information is usually presented in the Statement of Intent and/or Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the planned quality is deemed to be presented if at least there is information about planned dates or deadlines of activities in the output.</li> </ul>	
Timeliness variance disclosure for each output	<p>For each output, calculate if the variance between actual timeliness measures and planned/budgeted timeliness measures is 10% or more of the planned/budgeted timeliness measures. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for the variance. If there is an explanation, put a score of 1 in the appropriate cell. This information is usually presented in the Annual Report.</p> <ul style="list-style-type: none"> <li>For binary-type output (output with a quantity of one), the variance is determined if there is information about planned dates are not similar the actual dates of activities in the output or if there is information that deadlines are not met.</li> </ul>	
Actual cost measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on actual costs is presented. This information is usually presented in the Annual Report.</p>	
Planned cost measures for each output	<p>For each output, put a score of 1 in the appropriate cell if the information on planned/budgeted costs is presented. This information is usually presented in the Statement of Intent and/or Annual Report.</p>	
Cost variance disclosure for each output	<p>For each output, calculate if the variance between actual cost and planned/budgeted cost is 10% or more of the planned/budgeted cost. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for the variance. If there is an explanation,</p>	

<b>Items</b>	<b>Instructions</b>	<b>Control</b>
	put a score of 1 in the appropriate cell. This information is usually presented in the Annual Report.	

## SCORING INSTRUCTION FOR GENERAL OUTCOME INFORMATION

The scoring is applied to the Statement of Intent and the corresponding Annual Report. The coder should check both documents in following the instructions below.

<b>Items</b>	<b>Instruction</b>	<b>Control</b>
Description of goals and objective	If there is a section that discusses goals/objectives put a score of 1 in the appropriate cell.	
Link between goals or objectives and outcomes	If there is an assertion of the link between goals/objectives and outcomes, put a score of 1 in the appropriate cell.	
Elaborate explanation on the link between goals and objectives	If the above assertion is further explained, put a score of 1 in the appropriate cell. The elaborate explanation is deemed to be properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Description of cost effectiveness	If there is a short description on cost effectiveness of the interventions, put a score of 1 in the appropriate cell. The description is deemed to be presented if there is at least information on names/titles of cost effectiveness, a paragraph of discussion on cost effectiveness, and names/titles of measures used.	
Explanation on measures	If each measure is further explained, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Actual measures of cost effectiveness	If actual result(s) on measures of cost effectiveness is/are presented, put a score of 1 in the appropriate cell.	
Planned measures of cost effectiveness	If planned/targeted result(s) on measures of cost effectiveness is/are presented, put a score of 1 in the appropriate cell.	
Description of organisational health and capability	If there is a short description on organisational health and capability, put a score of 1 in the appropriate cell. The description is deemed to be presented if there is at least information on names/titles of organisational health and capability, a paragraph of discussion on organisational health and capability, and names/titles of measures used.	
Explanation on measures	If each measure is further explained, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Actual measures of organisational health	If actual result(s) on measures of organisational health and capability is/are presented, put a score of	

<b>Items</b>	<b>Instruction</b>	<b>Control</b>
and capability	1 in the appropriate cell.	
Planned measures of organisational health and capability	If planned/targeted result(s) on measures of organisational health and capability is/are presented, put a score of 1 in the appropriate cell.	

## SCORING INSTRUCTION FOR INDIVIDUAL OUTCOME INFORMATION

The scoring is applied to the Statement of Intent and the corresponding Annual Report. The coder should check both documents in following the instructions below.

<b>Items</b>	<b>Instruction</b>	<b>Control</b>
Identification of outcomes	Identify the outcomes and list them horizontally in the scoring sheet for output level. The outcomes identified are the highest-level ones.	
Title or short description of outcomes	For each outcome identified, if there is a short description on outcomes targeted, put a score of 1 in the appropriate cell. The description is deemed to be presented if there is at least information on names/titles of outcomes, a paragraph of discussion on outcomes, and names/titles of measures used.	
Explanation of outcomes	For each outcome identified, if each measure is further explained, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Outcome framework outlining hierarchies of outcomes	For each outcome identified, if there is an outcome framework that outlines the relationship/link between different levels of outcomes, put a score of 1 in the appropriate cell. The outcome framework for all outcomes may be presented in one tabular or graphical format. If this is the case, then the information is deemed to be disclosed.	
Explanation of the links in the outcome framework	For each outcome identified, if the links in the outcome framework is further explained, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Broad statement of outcome targets (qualitative targets)	For each outcome identified, if there is a statement on the outcomes targeted, put a score of 1 in the appropriate cell. If a description of quantitative outcome target is presented, then the information is deemed presented.	
Quantitative outcome targets	For each outcome identified, if there is/are quantitative outcome target(s), put a score of one in the appropriate cell.	
More than one outcome targets are presented	For each outcome identified, if there is more than one indicator for outcomes targeted, put a score of one in the appropriate cell.	



<b>Items</b>	<b>Instruction</b>	<b>Control</b>
A simple assertion of the link between outputs and outcomes	For each outcome identified, if there is an assertion of which outputs contribute to the outcome (intervention logic), put a score of 1 in the appropriate cell. The assertion may be presented in one tabular or graphical format. If this is the case, then the information is deemed to be disclosed.	
An elaborate explanation of the intervention logic	For each outcome identified, if the link between outputs and outcome (intervention logic) is further explained, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Discussion of output alternatives	For each output, if the presentation on intervention logic contains discussion on output alternatives that have been considered in the planning process, put a score of 1 in the appropriate cell. The description is deemed to be presented if there is at least information on names/titles of alternative outputs and a paragraph of discussion on the outputs.	
Identification of other factors/risks influencing outcomes	For each output, if the presentation on intervention logic contains discussion on factors/risks that are considered to influence the outcomes, put a score of 1 in the appropriate cell. The description is deemed to be presented if there is at least information on names/titles of alternative outputs and a sentence of description on the outputs.	
Explanation of the factors/risks influencing outcomes	For each outcome identified, if the factors/risks that influence outcome are elaborated, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least a paragraph of discussion is presented, with a minimum of five sentences.	
Management assessment or projection on the factors/risks	For each outcome identified, if there is an assessment or projection on the factors/risks that influence outcome, put a score of 1 in the appropriate cell. The explanation is deemed properly disclosed if at least there is a discussion on the current state and the projection of future state of the factors/risks.	
Planned/targeted outcomes (quantitative indicators)	For each outcome identified, if there is/are quantitative target(s) on outcome indicators presented, put a score of 1 in the appropriate cell.	

<b>Items</b>	<b>Instruction</b>	<b>Control</b>
Current results on outcome indicators (measures from the current period)	For each outcome identified, if there is a disclosure of actual outcome indicator(s) measured in the current period of the Statement of Intent and Annual Report, put a score of 1 in the appropriate cell.	
Past results on outcome indicators (measures from past periods)	For each outcome identified, if there is a disclosure of actual outcome indicator(s) measurement of previous periods, put a score of 1 in the appropriate cell.	
Discussion on outcome variance	For each outcome identified, calculate if the variance between actual outcome indicator(s) and planned/budgeted outcome indicator(s) is 10% or more of the planned/budgeted cost. If no, put n/a (not applicable) in the proper cell. If yes, identify whether there is an explanation for the variance. If there is an explanation, put a score of 1 in the appropriate cell. If planned and/or actual outcome indicator(s) is/are not available, put n/a (not applicable) in the appropriate cell	
Comparative outcome data	For each outcome identified, if there is a disclosure of similar outcome indicator(s) measurement from other entities and/or jurisdictions, put a score of 1 in the appropriate cell.	

## APPENDIX G

### Scoring Example

SCORING AT OUTPUT CLASS LEVEL FOR DEPARTMENT XYZ  
THE STATEMENT OF INTENT AND ANNUAL REPORT FOR 2003/2004

DISCLOSURE ITEMS		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	avg
A	Output Description											
1	Description of output class	1	1	1	0	1	1	0	1	0	1	0.7
2	A list of detailed outputs	1	1	1	0	1	1	0	1	1	1	0.8
3	Description of detailed outputs	1	1	1	0	1	1	1	1	1	1	0.9
4	Explanation of output class change	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
B	Cost disclosure											
5	Total actual cost for each output class	1	1	0	1	1	0	1	1	1	1	0.8
6	Total planned cost for each output class	1	1	0	1	1	1	0	1	1	1	0.8
7	Cost allocation method disclosure	0	0	0	0	0	0	0	0	0	0	0
8	Variance disclosure	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note:

The average score is used to calculate the total score of disclosure items. It represents the comprehensiveness of item information. In this example, the total score of information at output class level is 5 from 6 applicable items.

SCORING EXAMPLE  
 SCORING AT OUTPUT LEVEL FOR DEPARTMENT XYZ  
 THE STATEMENT OF INTENT AND ANNUAL REPORT FOR 2003/2004

DISCLOSURE ITEMS	D1.01	D1.02	D1.03	D1.04	D1.05	D2.01	D2.02	D2.03	D2.04	D2.05	D3.01	D4.01	D4.02	D8.01	D9.01	D9.02	D9.03	D10.01	D10.02	D10.03	avrg
C Presentation of quantity measures																					
9 Actual quantity	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0.95
10 Planned quantity	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0.95
11 Variance disclosure	n/a	n/a	n/a	n/a	1	n/a	n/a	n/a	n/a	n/a	1	n/a	n/a	1	n/a	1	n/a	n/a	n/a	n/a	1
D Quality description																					
12 List of relevant quality dimensions	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0.975
13 Descriptions of relevant quality dimensions	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0.375
E Quality benchmark used																					
14 Subjective (internal) benchmark	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0.55
15 External benchmark	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0.575
16 External assessment of external benchmark	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0.575
F Presentation of quality measures																					
17 Actual quality measures	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0.975
18 Planned quality measures	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0.975
19 Variance disclosure	1	n/a	n/a	n/a	n/a	n/a	n/a	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1	n/a	n/a	n/a	n/a	1
G Presentation of timeliness measures																					
20 Actual timeliness measure	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0.275
21 Planned timeliness measure	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0.275
22 Variance disclosure	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1	1	1	n/a	n/a	n/a	n/a	1
H Presentation of cost measures																					
23 Actual cost measure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24 Planned cost measure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25 Variance disclosure	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note:

Detailed outputs for output class D5-D7 are not skipped in the above table. The average score is used to calculate the total score of disclosure items. The average scores presented in the table are the scores for all detailed outputs for output class D1-D10. It represents the comprehensiveness of item information. In this example, the total score of information at output level is 10.45 from 17 applicable items. Thus, total score for output information is 15.45 out of 33 applicable items, resulting in 0.67 in terms of relative score.

SCORING EXAMPLE  
 SCORING FOR OUTCOME INFORMATION FOR DEPARTMENT XYZ  
 THE STATEMENT OF INTENT AND ANNUAL REPORT FOR 2003/2004

DISCLOSURE ITEMS		Obj 1	Avg
<b>A Goals/objectives identification</b>			
1	Description of goals/objectives	1	1
2	Link between goals/objectives and outcomes in graphical/tabular form	1	1
3	Elaborate explanation of the links between goals/objectives and outcomes	0	0
<b>H Cost effectiveness of the interventions</b>			
22	Descriptions of cost effectiveness	0	0
23	Explanation of measures of cost effectiveness	0	0
24	Actual measures	0	0
25	Planned measures	0	0
<b>I Organisation health and capability to deliver outcomes</b>			
26	Descriptions of organisation health and capability	1	0
27	Explanation of measures of organisation health and capability	0	0
28	Actual measures	0	0
29	Planned measures	0	0

DISCLOSURE ITEMS		OC1	OC2	OC3	Avg
<b>B Outcome description</b>					
4	Title or short description of outcomes	1	1	1	1
5	Explanation of each outcomes	1	1	1	1
6	Outcome framework outlining high-level and low-level outcomes	1	1	1	1
7	Explanation of the links between high-level and low-level outcomes	0	0	0	0
<b>C Outcome target</b>					
8	Broad statement of targets	1	1	1	1
9	Quantitative targets	1	1	1	1
10	Multi-dimension targets	1	1	1	1
<b>D Outcome-output link discussion</b>					
11	Simple linking statement (tabular or graphical form)	1	1	1	1
12	Explanations to justify the links (intervention logic)	0	0	0	0
13	Discussion of output alternative	0	0	0	0
<b>E Discussion of other factors influencing outcomes</b>					
14	Identification of the factors	0	0	0	0
15	Explanation of the factors	0	0	0	0
16	Management assessment of projections of the factors	0	0	0	0
<b>F Presentation of outcome measures</b>					
17	Planned outcomes	0	0	0	0
18	Variance discussion	n/a	n/a	n/a	n/a
19	Past outcomes (measures from previous periods)	1	1	1	1
20	Current outcomes (measures from current periods)	1	1	1	1
21	Comparative outcome data from other organisations and/or countries	0	0	0	0

Note:

The average score is used to calculate the total score of disclosure items. It represents the comprehensiveness of item information. Thus, total score for outcome information is 9 out of 28 applicable items, resulting in 0.32 in terms of relative score.

# APPENDIX H

## Ethics Information Sheets

### ETHICS INFORMATION SHEET FOR PARTICIPANTS FROM CENTRAL AGENCIES AND THE OFFICE OF AUDITOR GENERAL

VICTORIA UNIVERSITY OF WELLINGTON  
*Te Whare Wananga o te Upoko o te Ika a Maui*



#### Interviewee Information Sheet for a Study of Performance Reporting by New Zealand Central Government Agencies

**Researcher:** Rahmadi Murwanto: School of Accounting and Commercial Law, Victoria University of Wellington

I am a PhD student in the School of Accounting and Commercial Law at Victoria University of Wellington. As part of this degree I am undertaking a research project leading to a thesis. The project I am undertaking is assessing longitudinally the increase in the quality of performance information as a result of the Managing for Outcome Initiative. I will also identify the internal and external factors to influence the increase in the quality of performance information over time. The University required that ethics approval be obtained for research involving human participants. The ethics approval has been obtained.

I am inviting officials in the New Zealand Central Government Agencies to participate in this study. Participants will be interviewed on a set of open questions on their view of the development of Managing for Outcome Initiative and/or its impacts on the quality of their institution's performance information. Each interview will last around 60-90 minutes. The interview will be at the place of your convenience. It will be recorded and transcribed. Once the interview is transcribed, you have the right to check the full interview transcript to clarify or to change your response if you wish to no later than 5 April 2010. A follow-on interview may be conducted to further clarify the responses before 5 April 2010.

Should any participants feel the need to withdraw from the project, they may do so without question at any time before 5 April 2010 when the data is analysed. Just let me know at any time before this date.

Responses collected will form the basis of my research project and will be put into a written report on an anonymous basis and will thus not be associated with names of participants. However, it may be possible that your organisations to be identified when specific functions associated with your responses are discussed in the report. All material collected will be kept confidential. No other person besides me and my supervisors, Prof Tony van Zijl and Dr Bhagwan Khanna, will see the records and transcripts. The thesis will be submitted for marking to the School of Accounting and Commercial Law and deposited in the University Library. It is intended that one or

more articles will be submitted for publication in scholarly journals. The records and transcripts will be destroyed two years after the end of the project.

If you have any questions or would like to receive further information about the project, please contact me at 021 264 7312 or my supervisor, Prof Tony van Zijl, at the School of Accounting and Commercial Law at Victoria University, P O Box 600, Wellington, phone 04 463 5329, and Dr Bhagwan Khanna at the School of Accounting and Commercial Law at Victoria University, P O Box 600, Wellington, phone 04 463 5843.

Rahmadi Murwanto

Signed:

## **ETHICS INFORMATION SHEET FOR PARTICIPANTS FROM MINISTRIES/DEPARTMENTS DESIGNATED AS PREPARERS**

VICTORIA UNIVERSITY OF WELLINGTON  
*Te Whare Wananga o te Upoko o te Ika a Maui*



### **Interviewee Information Sheet for a Study of Performance Reporting by New Zealand Central Government Agencies**

**Researcher:** Rahmadi Murwanto: School of Accounting and Commercial Law, Victoria University of Wellington

I am a PhD student in the School of Accounting and Commercial Law at Victoria University of Wellington. As part of this degree I am undertaking a research project leading to a thesis. The project I am undertaking is assessing longitudinally the increase in the quality of performance information as a result of the Managing for Outcome Initiative. I will also identify the internal and external factors to influence the increase in the quality of performance information over time. The University required that ethics approval be obtained for research involving human participants. The ethics approval has been obtained.

I am inviting officials in the New Zealand Central Government Agencies to participate in this study. Participants will be interviewed on a set of open questions on their view of the development of Managing for Outcome Initiative and/or its impacts on the quality of their institution's performance information. Each interview will last around 60-90 minutes. The interview will be at the place of your convenience. It will be recorded and transcribed. Once the interview is transcribed, you have the right to check the full interview transcript to clarify or to change your response if you wish to no later than 5 April 2010. A follow-on interview may be conducted to further clarify the responses before 5 April 2010.

Should any participants feel the need to withdraw from the project, they may do so without question at any time before 5 April 2010 when the data is analysed. Just let me know at any time before this date.

Responses collected will form the basis of my research project and will be put into a written report on an anonymous basis and will thus not be associated with names and organisations of participants. All material collected will be kept confidential. No other person besides me and my supervisors, Prof Tony van Zijl and Dr Bhagwan Khanna, will see the records and transcripts. The thesis will be submitted for marking to the School of Accounting and Commercial Law and deposited in the University Library. It is intended that one or more articles will be submitted for publication in scholarly journals. The records and transcripts will be destroyed two years after the end of the project.

If you have any questions or would like to receive further information about the project, please contact me at 021 264 7312 or my supervisor, Prof Tony van Zijl, at the School of Accounting and Commercial Law at Victoria University, P O Box 600, Wellington,



phone 04 463 5329, and Dr Bhagwan Khanna at the School of Accounting and Commercial Law at Victoria University, P O Box 600, Wellington, phone 04 463 5843.

Rahmadi Murwanto

Signed:



## CONSENT TO PARTICIPATION IN RESEARCH

**Title of project: Performance Reporting by New Zealand Central Government Agencies**

**Please tick the following statement to indicate your consent to participation in this research.**

I have been given and have understood an explanation of this research project. I have had an opportunity to ask questions and have them answered to my satisfaction.

Y  N

I wish to check the interview transcript before 5 April 2010 (before data collection and analysis is complete).

Y  N

I understand that a follow on interview may be conducted to further clarify my responses before 5 April 2010 (before data collection and analysis is complete).

Y  N

I understand that I may withdraw myself (or any information I have provided) from this project and the tape recordings will be wiped no later than 5 April 2010 (before data collection and analysis is complete) without having to give reasons.

Y  N

I understand that any information I provide will be kept confidential to the researcher and the supervisor, the published results will not use my name, and that no opinions will be attributed to me in any way that will identify me.

Y  N

I understand that the tape recording of interviews will be electronically wiped on 31 August 2012 (two years after the project ends) unless I indicate that I would like them returned to me.

Y  N

**Name:**

**Signed:**

## **Interview Research Instrument**

### **INTERVIEW TARGETS AND THE LIST OF QUESTIONS**

#### **Top management (CEO or CFO) of selected NZ central government agencies**

1. What are the expectations of the Minister, Parliament, and/or other parties (such as the Office of Auditor General, the Treasury and the State Services Commissions) on the level of the quality of performance information reported?
2. Does the quality of non-financial performance information (output and outcome information) reflect your efforts in informing the stakeholders on your achievement in managing the agency? Why or Why not?
3. Are there any incentives or benefits personally and organisation-wise that might be received or are received from reporting the performance information (output and outcome information) in your corporate report? If so, what are the incentives?
4. Are there any impacts your reputation as a professional executive of a public agency by reporting the performance information in the corporate reports? If so, what are the impacts?
5. Is there any concern for revealing more performance information? What are the concerns, if any?
6. How do you create a boundary of the level of performance information to be disclosed to the public?
7. Do you have any concern that any reported information might be used against you or your agency? What are the concerns, if any?
8. Are there any constraints in reporting your performance? If so, what are the constraints that you identify?
9. What is your opinion on the reporting standards or reporting requirement for output and outcome information? Are they helpful in prescribing you what to report?
10. Do you consider that the agency's reported performance (its outputs and outcomes) is influenced by some factors beyond the agency's control?
11. What is your opinion if some parts of your compensation (salary and/or bonus) are based on the reported figure of the organisation's output and/or outcome?

#### **Staff in charge of the SOI and Annual Report preparation of selected NZ central government agencies**

1. How output measures in the SOI and the Annual Report is developed? Does it involve these parties below for the development of output information? If so, what are their roles?
  - a. The Ministers
  - b. Members of Parliament
  - c. Users of your products/services
  - d. Auditors
  - e. Others (please identify)

2. The TPA-9 recommends for 5 components of output information (quantity, quality, timeliness, location and cost). Why are some components not reported (such as quality, location and/or cost per output)?
3. How outcome measures in the SOI and the Annual Report is developed? Does it involve these parties below for the development of output information? If so, what are their roles?
  - a. The Ministers
  - b. Members of Parliament
  - c. Users of your products/services
  - d. Auditors
  - e. Others (such as the Treasury, the SSC and the Office of Auditor General)
4. Are the outcomes of your agency easily identified, developed and agreed by the parties involved in the development of outcome measures? Are there any disagreements during the process? If so, how do you resolve the disagreement?
5. Considering that your agency delivers multiple services, do you consider that there are some difficulties in developing a coherent outcome framework for your agency?
6. What are the difficulties in developing the outcome framework?
7. How the performance information is collected, summarised, and reported?
8. Are there specific resources used in these processes?
9. How the capability to provide performance information is developed? Is there any specific information system and/or dedicated staff for this? Is there any specific training for the staff responsible for the preparation of performance information reported?
10. Do you get any supports from these agencies below in the process of the development of the capability and/or the development of performance measures? If so, what are the supports?
  - a. The Ministers
  - b. Member of Parliament
  - c. The Treasury
  - d. The SSC
  - e. The Office of Auditor General
  - f. Other institutions (please identify)

**The Treasury: officials involved in the MfO initiative. The list of the questions:**

1. How do you involve in the MfO Initiative? Describe your involvement historically in the initiative.
2. What is Pathfinder Project? What is its relation with MfO Initiative? What is the Treasury role in the Pathfinder Project?
3. What are the impacts of MfO Initiative in the performance reporting practice in NZ public sector? Are these impacts are expected? Does the current reporting practice meet the targeted impacts set in the MfO Initiative?
4. What are the important roles delivered by the Treasury in the MfO initiative?
5. How the guidance documents for performance reporting (guidance for MfO, guidance for SOI preparation and guidance for Annual Report preparation) are developed?
6. Did the Treasury conduct any evaluation processes during the implementation of MfO initiative, especially related to the quality of performance information reported? If so, what are major issues identified from this process?
7. Are there any roles adopted or other actions taken by the Treasury to improve the quality of performance information of performance information?

8. Is there any specific support from the Treasury to each government agency in its preparation of performance information?
9. In the budget allocation decision, is performance information is used as an input?

**State Services Commission: officials involved in the MfO initiative. The list of the questions:**

1. How do you involve in the MfO Initiative? Describe your involvement historically in the initiative.
2. What is Pathfinder Project? What is its relation with MfO Initiative? What is the SSC role in the Pathfinder Project?
3. What are the impacts of MfO Initiative in the performance reporting practice in NZ public sector? Are these impacts are expected? Does the current reporting practice meet the targeted impacts set in the MfO Initiative?
4. What are the important roles delivered by the SSC in the MfO initiative?
5. Does the SSC involve in the development of guidance documents for performance reporting? If so, what are the SSC's roles?
6. Did the SSC conduct any evaluation processes during the implementation of MfO initiative, especially related to the quality of performance information reported? If so, what are major issues identified from this process?
7. Are there any roles adopted or other actions taken by the SSC to improve the quality of performance information of performance information?
8. Is there any specific support from the Treasury to each government agency in its preparation of performance information?
9. In the evaluation of CEO' performance and/or determination of their compensation, is performance information is used as an input?

**The Office of Auditor General: officials involved in the MfO initiative or performance reporting audit policy and the auditors. The list of the questions:**

**For officials involved in the MfO initiative or performance reporting audit policy:**

1. How do you involve in the MfO Initiative? Describe your involvement historically in the initiative.
2. What are the important roles delivered by the OAG in the MfO initiative?
3. Are there any roles adopted or other actions taken by the OAG to improve the quality of performance information of performance information?
4. Is there any specific support from the OAG to each government agency in its preparation of performance information?

**The auditors:**

1. Do you provide a review of outcome information reported in the SOI and the Annual Report of the Agency? How the review is done?
2. How the audit is done for the Statement of Service Performance?
3. What reporting standards/guidance do you use in the review of the SOI and the audit of the Statement of Service Performance?
4. How much time and resources (in a percentage of total time and resources spent for the audit) do you spend for auditing/reviewing the Statement of Service Performance and the disclosure of outcome information?
5. Do you provide any specific advices to improve the outcome information reported? The output information reported?

This page is blank.