

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# **Values-Based Evaluative Management**

**An integrated and adaptive approach to enhance  
inclusion, development effectiveness, governance, and  
sustainability**

**A thesis presented in partial fulfilment of the  
requirements for the degree of**

**Doctor of Philosophy  
in  
Development Studies**

**at Massey University, Palmerston North,  
New Zealand.**

**Catherine Mary Averill**

**2021**



# Abstract

Sustainable development, an articulated goal of development practice in the 21st century (United Nations Development Programme, 2015a) now needs to be inclusive, based on multi-level systems of accountability, and have robust governance. This thesis proposes alternative *evaluative management* values and principles for inclusive sustainable development that are values-based, integrated, and adaptive. It suggests the way such values and high-level principles could underpin and reposition development, management, and evaluation approaches.

An initial idea behind this research was that there needed to be a better way to connect strategic evaluative approaches within management and potentially the new sustainable development goals in international and national development. The impacts and significance of changes for both the broader development context and governance systems of country-level development, and the management and evaluation practices, were examined in the context of countries and donors in two Pacific settings: Papua New Guinea and Aotearoa/New Zealand.

To achieve the sustainable development goals, current management theory and practices needs to be reconsidered. This research pointed to the emergence of *evaluative management* as an identifiable theoretical and instrumental discourse and knowledge frame repositioning and integrating existing management discourses underpinned by values and principles relating to strategic planning, performance management and governance. This thesis proposes that a new model of integrated management – called, in this context, *evaluative management* that is premised on three values (inclusion, partnership and participation) and three high-level principles (relationality, contextual sensitivity and adaptive response), is needed to underpin such considerations.

The potential of *evaluative management* can only be realised if it is enacted through values and principles that are well communicated and widely understood. This may include effective interaction and communication between different levels of governments, agencies, sectors, regions and communities including non-governmental organisations, private sector, and development partners. This new model of integrated management would also provide the capacity to address inclusion, governance, accountability, and sustainable development with more effective strategic evaluative practices.

# Acknowledgements

To the people from Aotearoa New Zealand and Papua New Guinea who contributed to this research, thank you for your time, support, and contribution. I am so grateful and appreciative of the managers and evaluators who shared their professional and personal views as part of this research. I hope I have done justice to your inputs, reflections, and insightful contributions.

To my supervisors – Robin Peace and Glenn Banks – I sincerely thank you for challenging me to think more critically and encouraging me as I worked to make sense of the complexity of the emergent findings. My appreciation is more than an acknowledgement. This research became a humbling journey for me as a practitioner. I had a hunch that there was more to exploring and thinking more deeply, to learning from my findings and navigating a way to write about a complex topic. I hope that people can reflect on and share what this research means, what is happening already, ways forward, and potentially how this research can contribute to knowledge and practice.

My research drew on my previous academic studies of an undergraduate science degree majoring in geography (human and physical), a Post-Graduate Diploma of Secondary Teaching, a Master of Business Administration, and a Post-Graduate Diploma in Social Sector Evaluation Research. My experience as an educator, manager, and evaluator contributed to this research, particularly in making sense of existing and emergent concepts.

To my evaluation colleagues, thank you for your support and encouragement.

To my black lab, George – thank you for sharing some of this PhD journey, sitting patiently under my desk until you would remind me it was time for a walk.

To my parents Colin and Valerie, whose support meant that the writing up of this research was possible to complete.

To Māori – *tangata whenua* (indigenous peoples) of Aotearoa New Zealand – and Pacific peoples and other cultures, I hope that sharing this research may contribute to a more inclusive way forward.

To families from the mosque attack on 15 March 2019 in Christchurch – to living in a world with shared values where people are respected, and differences and similarities celebrated and nurtured.

# Table of Contents

<b>ABSTRACT .....</b>	<b>III</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>IV</b>
<b>LIST OF TABLES.....</b>	<b>VII</b>
<b>GLOSSARY .....</b>	<b>I</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>IV</b>
<b>PART A: FOREGROUND MATERIAL.....</b>	<b>1</b>
CHAPTER 1: INTRODUCTION .....	1
1.1 Positionality .....	3
1.2 Background .....	4
1.3 Research context.....	6
1.4 Context of the study .....	7
1.5 Conceptual research frame.....	8
1.6 Research questions .....	10
1.7 Problem statement .....	11
1.8 Research timeline and thesis structure .....	13
CHAPTER 2: RESEARCH KNOWLEDGE CONTEXT .....	15
2.1 Approach to the literature analysis.....	16
2.2 Terminology.....	19
2.3 Development .....	20
2.4 Management and accountability.....	22
2.5 Results-based management (RBM) .....	23
2.6 Country systems .....	28
2.7 Impact model.....	35
2.8 Themes .....	37
2.9 Knowledge space for this research.....	40
CHAPTER 3: METHODOLOGY .....	42
3.1 Theoretical perspective and research frame .....	42
3.2 Research questions .....	47
3.3 Research methodology.....	48
3.4 Research phases and methods.....	50
3.5 Data analysis.....	58
3.6 Research method limitations .....	60
3.7 Overview of the evolution of the research .....	62
3.8 Overview of this research.....	64
CHAPTER 4: PRACTICE CONTEXTS .....	65
4.1 Aotearoa/New Zealand country governance and management .....	65
4.2 Performance management and evaluation.....	71
4.3 Papua New Guinea .....	74
4.4 Case research context.....	78
<b>PART B: PRACTICE WITHIN THE CONTEXTUAL LAYERS.....</b>	<b>81</b>
CHAPTER 5: IDEAS, INTERVENTIONS AND FRAMEWORKS LAYER.....	81
5.1 Aotearoa/New Zealand document review .....	81

5.2	Papua New Guinea document review .....	85
5.3	Rating of impact model dimensions – ideas layer .....	87
5.4	Substantive progress – ideas layer .....	88
5.5	Substantive concepts requiring more attention .....	95
5.6	Emergent concepts requiring inclusion .....	113
5.7	Chapter summary .....	122
CHAPTER 6: INSTITUTION LAYER .....		123
6.1	Rating of impact model dimensions – institution layer .....	123
6.2	Substantive progress – institution layer .....	125
6.3	Substantive concepts requiring more attention .....	126
6.4	Emergent concepts requiring inclusion .....	157
6.5	Chapter summary .....	161
CHAPTER 7: WIDER INFRASTRUCTURE LAYER .....		162
7.1	Rating of impact model dimensions – wider infrastructure layer .....	162
7.2	Substantive progress – Wider infrastructure layer .....	163
7.3	Substantive concepts requiring more attention .....	171
7.4	Emergent concepts – requiring inclusion .....	189
7.5	Chapter summary .....	196
<b>PART C: DISCUSSION AND CONCLUSIONS .....</b>		<b>199</b>
CHAPTER 8: CONCEPTS LINKING FRAMEWORKS, MANAGEMENT SYSTEMS AND PRINCIPLES .....		199
8.1	Critical discussion .....	200
8.2	Potential reconfiguring for the role of strategic evaluation .....	218
8.3	Identifiable concepts requiring more attention and inclusion .....	220
8.4	High-level principles .....	233
8.5	Chapter summary .....	236
CHAPTER 9: VALUES-BASED EVALUATIVE MANAGEMENT .....		237
9.1	A new model .....	239
9.2	Good practice concepts .....	243
9.3	Chapter summary .....	245
CHAPTER 10: CONCLUSIONS AND IMPLICATIONS .....		247
10.1	Conclusions .....	247
10.2	Implications .....	253
<b>BIBLIOGRAPHY .....</b>		<b>257</b>
<b>APPENDICES .....</b>		<b>287</b>
APPENDIX A: LITERATURE REVIEW CONCEPT MAP (2011) .....		288
APPENDIX B: ALIGNING CONCEPTS AT DIFFERENT LEVELS .....		289
APPENDIX C: ETHICS APPLICATION .....		290
APPENDIX D: DATA COLLECTION TOOLS .....		293
APPENDIX E: SAMPLE OF QUALITATIVE FINDINGS ANALYSIS .....		305
APPENDIX F: QUALITATIVE RESEARCH FINDINGS THEMATIC SUMMARY .....		325
APPENDIX G: FULL SUMMARY TABLE OF IDENTIFIED CONCEPTS FROM RESEARCH .....		333
APPENDIX H: LIST OF ORGANISATIONS WHO PARTICIPATED IN THIS RESEARCH .....		342

## List of tables

Table 1. A/NZ agency rating data.....	82
Table 2. Rating table for impact model dimensions, A/NZ case study results.....	87
Table 3. Substantive progress: ideas, interventions and frameworks layer.....	88
Table 4. Substantive concept requiring more attention: ideas, interventions and frameworks layer.....	95
Table 5. Emergent concepts requiring inclusion: <i>ideas, interventions and frameworks</i> layer.....	113
Table 6. Rating table for impact model dimensions: institution layer, A/NZ case study results.....	124
Table 7. Substantive progress: institution layer.....	125
Table 8. Substantive conceptual areas requiring more attention: institution layer.....	127
Table 9. Emergent concepts requiring inclusion: institution layer.....	157
Table 10. Rating table for impact model dimensions: wider infrastructure layer, A/NZ case study results.....	162
Table 11. Substantive progress: wider infrastructure layer.....	163
Table 12. Substantive conceptual areas requiring increased attention: wider infrastructure layer.....	171
Table 13. Emergent conceptual areas requiring inclusion: wider infrastructure layer.....	189
Table 14. Key identifiable good practice concepts requiring further attention or inclusion.....	221
Table 15. Values-Based Evaluative Management: guiding values, settings, high-level principles and associated good practice concepts.....	241
Table 16. Good practice concept explanation under a Country Planning and Evaluative Management System approach.....	243
Table 17. Analysis template for public sector agencies/sectors.....	299
Table 18. Concepts identified from the research.....	333



# List of Figures

<b>Figure 1.</b> Research timeline and thesis structure. ....	13
<b>Figure 2.</b> Initial research impact model. ....	36
<b>Figure 3.</b> Ontological conception of social reality (Bhaskar, 1979/1998, p. 51). ....	42
<b>Figure 4.</b> Pawson’s hierarchy of contextual layers showing “the intervention is the product of its context” (2006, p. 31). ....	43
<b>Figure 5.</b> Emergent conceptual research typology and primary analytical frame....	44
<b>Figure 6.</b> Layers of reality and the role of abduction (Wuisman, 2005, p. 384). ....	45
<b>Figure 7.</b> Updated impact model – initial research frame (superimposed as an emergent research frame. ....	51
<b>Figure 8.</b> The typology of the country case study research and associated research methods.....	55
<b>Figure 9.</b> Structure of country case study object for A/NZ case study research...	56
<b>Figure 10.</b> Emergent code frames in NVivo. ....	59
<b>Figure 11.</b> Key accountability relationships in the A/NZ Public Service (Gill, 2008, p. 40).....	68
<b>Figure 12.</b> PNG National Planning Framework 2016 (DNPM, 2016, p.3).....	76
<b>Figure 13.</b> Country Planning and Evaluative Management System Model. ....	238

# Glossary

In this glossary, key terms used in this thesis are listed and the meanings of specific words relevant to this thesis topic are provided. The terms defined in this glossary were predominantly based on the *Glossary of Key Terms in Evaluation and Results-based Management* (Organisation for Economic Co-operation and Development [OECD], 2010c) unless specified or noted where the meanings of terms are widely accepted.

Accountability	"Obligation to demonstrate that work has been conducted in compliance with agreed rules and standards or to report fairly and accurately on performance results vis a vis mandated roles and/or plans" (p. 15)
Aotearoa	Te Reo (Māori language) word for New Zealand
Audit	"An independent, objective assurance activity designed to add value and improve an organization's operations" (p. 17)
Beneficiary	"The individuals, groups, or organizations, whether targeted or not, that benefit, directly or indirectly, from the development intervention" (p.18)
Country system	A country system is defined as "national arrangements and procedures for public financial management, procurement, audit, monitoring and evaluation, and social and environmental procedures" (OECD, 2010b, p.45).
Data collection tools	"Methodologies used to identify information sources and collect information during an evaluation" (p. 19)
Effectiveness	"The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance. Note: Also used as an aggregate measure of (or judgment about) the merit or worth of an activity, i.e. the extent to which an intervention has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable fashion and with a positive institutional development impact" (pp. 20-21)
Efficiency	"A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results" (p. 21)
Evaluation	"The systematic and objective assessment of an ongoing or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability" (pp. 21-22)
Feedback	"The transmission of findings generated through the evaluation process to parties for whom it is relevant and useful so as to facilitate learning. This may involve the collection and dissemination of findings, conclusions, recommendations and lessons from experience" (p. 23)
Events	"Events or outcomes are what critical realists investigate, that is the external and visible behaviours of people, systems and things as they occur, or as they have happened" (Easton, 2010, p. 120)
Goal	"The higher-order objective to which a development intervention is intended to contribute" (p. 24)
Hapū	Māori sub-tribe (Te Reo)
Impacts	"Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended" (p.24)
Indicator	"Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected

	to an intervention, or to help assess the performance of a development actor" (p.25)
Inputs	The financial, human, and material resources used for the development intervention
Iwi	Māori tribe (Te Reo)
Logical framework (Logframe)	"Management tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs, outcomes, impact) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure. It thus facilitates planning, execution and evaluation of a development intervention" (p. 27)
Kura	Māori Immersion school (Te Reo)
Mechanism	Mechanisms are "nothing other than the ways of acting of things"(Bhaskar, 1978. p. 14)
Monitoring	"A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with OECD indications of the extent of progress and achievement of objectives and progress in the use of allocated funds" (p 28)
Output	"The products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes" (p. 28).
Outcome	"The likely or achieved short-term and medium-term effects of an intervention's outputs" (p. 28)
Outcomes framework	An outcomes framework shows the hierarchy of key outcomes for a sector or overarching multi-programme (Duignan, 2004). This may include multiple outcome layers - sector, region, agency, and programme
Pākehā	European (A/NZ)
Policy	A course or principle of action adopted or proposed by an organisation or individual
Program logic	Diagram showing the links between inputs-outputs-outcomes/impacts, context and assumptions. This diagram is also often referred to as an intervention logic model.
Partners	"The individuals and/or organizations that collaborate to achieve mutually agreed upon objectives. Note: The concept of partnership connotes shared goals, common responsibility for outcomes, distinct accountabilities, and reciprocal obligations. Partners may include governments, civil society, non-governmental organizations, universities, professional and business associations, multilateral organizations, private companies" (p. 28)
Program theory	Describes the theories and rationale underpinning the intended changes resulting from activities
Results	"The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention" (p. 33)
Results framework	"The program logic that explains how the development objective is to be achieved, including causal relationships and underlying assumptions" (p.33)
Results-Based Management (RBM)	"A management strategy focusing on performance and achievement of outputs, outcomes and impacts" (p. 34)
Sector program evaluation	"Evaluation of a cluster of development interventions in a sector within one country or across countries, all of which contribute to the achievement of a specific development goal. Note: a sector includes development activities commonly grouped together for the purpose of public action such as health, education, agriculture, transport" (p. 35)

Stakeholders	Agencies, organisations, groups, or individuals who have a direct or indirect interest in the development intervention or its evaluation
Tangata whenua	Indigenous peoples (e Reo)
Theory of change	The theory of change is a description of the intended changes from an intervention.at a country, sector or programme level
Triangulation	“The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment. Note: by combining multiple data sources, methods, analyses or theories, evaluators seek to overcome the bias that comes from single informants, single methods, single observer, or single theory studies” (p.37)
Te Reo	Māori language
Whānau	Family (Te Reo)
Validity	“The extent to which the data collection strategies and instruments measure what they purport to measure” (p. 37)

## **List of abbreviations**

AG4	Accounting Standard Four (A/NZ)
A/NZ	Aotearoa//New Zealand
A/NZM	A/NZ Manager (research participant)
A/NZE	A/NZ Evaluator (research participant)
AR	Accountability Document Review
AusAID	Australian Aid Agency (latterly referred to as DFAT)
CACC	Central Agencies Coordinating Committee (PNG)
CE	Chief Executive
DFAT	Department of Foreign Affairs and Trade (Australia)
DHB	District Health Board (A/NZ)
DNPM	Department of National Planning and Monitoring (PNG)
DPLGA	Department of Provincial and Local Government Affairs (PNG)
DPMC	Department of Prime Minister and Cabinet (A/NZ)
GAO	Government Accountability Office (United States)
GoPNG	Government of Papua New Guinea
LFA	Logical Framework Approach
IRD	Inland Revenue Department (A/NZ)
LLG	Local-level Government
MAF	Ministry of Agriculture and Forestry (A/NZ)
MDG	Mid-term Development Goals
MFAT	Ministry of Foreign Affairs and Trade
MfDR	Managing for Development Results
MTDP	Mid-term Development Plan (PNG)
M&E	Monitoring and Evaluation
MSD	Ministry of Social Development (A/NZ)
NGO	Non-Governmental Organisation

OAG	Office of Auditor General( A/NZ)
OECD	Organisation for Economic Co-operation and Development
OECD-DAC	Organisation for Economic Co-operation and Development – Development Assistance Committee
PIF	Performance Improvement Framework (A/NZ)
PNG	Papua New Guinea
PNGDP	Papua New Guinea Development Partner (research participant)
PNGE	Papua New Guinea Evaluator (research participant)
PNGM	Papua New Guinea Manager (research participant)
QF	Qualitative findings
RBM	Results-Based Management
RQ	Research question
RT	Rating table
SDGs	Sustainable Development Goals
SOI	Statement of Intent (A/NZ)
SSC	SSC(A/NZ)
StaRS	Strategy for Responsible Sustainable Development (PNG)
TPK	Te Puni Kōkiri (A/NZ Māori Development Agency)
UN	United Nations
UNDP	United Nations Development Programme



# Part A: Foreground material

## Chapter 1: Introduction

The growing significance placed on more sustainable forms of development has brought into focus the connections between “economic activities, state policies, natural resources, the ecosystem and how societies function” (Lanoska, 2018, p. 177). This shift away from a primary emphasis on economic growth and progressive industrialisation requires different ways of appraising and responding to changing social, environmental, and economic conditions. As a senior evaluator working with managers and evaluators in two different development contexts – the law and justice sector in Papua New Guinea (PNG) and the transport sector in Aotearoa/New Zealand (A/NZ) – I encountered challenges such as:

- how to link strategy and evaluation knowledge and practice concepts
- when to involve different stakeholder perspectives and inputs
- how to work effectively with managers and evaluators across different roles in different settings
- what forms of evaluation to use to assess development effectiveness and identify aspects for improvement?

This research describes a journey towards a more comprehensive understanding of how and in what ways evaluation may be able to play a more strategic role in expanding development practice. My research included a focus on results-based frameworks and a set of values, principles, and informed concepts and practices that may better integrate evaluation into development systems and governance. My thesis initially appropriated international development agency ideas and understandings of development, such as those of the Organisation for Economic Co-operation [OECD], 2005a), rather than engaging critically with the notion of development as displayed by Fforde (2017). My research findings concurred with Horner and Hume’s (2019) more recent view of development with a global “converging divergence” (p. 496) of increasing inequalities within and between countries. They cited Raworth (2018) who outlined “we’re all developing countries now” (p. 504) with issues of climate change, inequality and sustainability of development outcomes.

From 2000 onwards, in international development contexts, evaluation theories, approaches, and practices were influenced by a shift from stand-alone development programmes, such as the *Family Health and Rural Improvement Program* in the Tari



Province in PNG (Vail, 2007), to country, sector, and agency focused examples (Conlin & Stirrat, 2008). Changes in development, management, and evaluation approaches and practices were needed in response to an increased emphasis on development at country, sector, and agency levels rather than at programme or project levels (Dahler-Larsen, 2012; Patton, 2011; Rist, 2006). An emerging trend from 2008 onwards was the growth of more coordinated country, sector, and agency planning approaches within diverse national development contexts. Chimhowu, Hume, and Munro (2019) noted that by 2018 over 80 percent of the world's population lived in countries with national development plans. They considered this expanding trend in 'new' national development planning may have implications for countries' capacity and commitments to be able to deliver on their plans effectively.

Increasingly, governance has become a growing focus in development discourse this century (Betts & Wedgewood, 2011; Dahler-Larsen & Boodhoo, 2019; Williams & Hummelbrunner, 2011). Grindle (2007), in referring to research undertaken by Kaufmann, Kraay, and Mastruzzi (2002), emphasised that good governance was necessary for development as it may enhance inclusion and development effectiveness. My thesis was that current evaluative approaches in the management and governance of development, particularly with the Sustainable Development Goals (SDGs) (United Nations Development Programme [UNDP], 2015b) may be insufficient to ensure robust, inclusive, and effective results. The overarching research question that I focused on to explore this argument was: *What good practice evaluative principles and concepts may enhance country, sector and agency development, management and governance knowledge and practice in different contexts?*

While development contexts were widely acknowledged to prevail in developing countries, this thesis argued that such normative aspirations also govern countries regarded as developed. I examined the idea that what may work in terms of understanding the value and impact of development in somewhere like PNG, a developing country with national and development partners, may also work in Aotearoa New Zealand (A/NZ), a developed country incorporating indigenous and migrant populations. Both these countries shared a relatively recent history and legacy of colonisation. PNG was automatically regarded as a site for development; A/NZ less so, given a history of predominant Western development perspectives displayed by consecutive governments. In the A/NZ context, national development focuses on separate government agencies and concepts for economic and social development (such as Māori and Pacific). Whereas in PNG, development planning is undertaken by central agencies using a country, sector, and provincial focus, with international

development aid funding being increasingly aligned to meet country and sector goals. A central thread of my argument comprised examining development, management, and evaluation concepts in these two constituencies to enhance development effectiveness and governance knowledge and practices.

## **1.1 Positionality**

I came to this PhD research as an evaluation practitioner with twelve years of experience in the field. In particular, I had worked in international development and in government roles. This experience shaped a pragmatic and largely uncritical view of how evaluation and policy operated at country and sector levels but did generate the beginnings of some questions that were not easily answered in the field.

I entered the practice of evaluation after undertaking management studies and educational planning and manager roles. I was particularly interested in how evaluation could align with strategic planning to support and measure progress and results in different country, sector and institution contexts (such as government agencies, non-governmental organisations (NGO) and private sector). In my practice, I often reflected on what I value (such as economic, social and environmental dimensions), drawing on my geographic academic background, when measuring progress and results.

My entry point into this interdisciplinary management and evaluation research in different national development contexts was based on my interest in the differences in stakeholder values within and between country management and evaluation practice contexts. When I first started as an evaluator, I was very accepting of common practice concepts such as the use of evaluation criteria and that different perspectives could be included. However, while working internationally, I encountered differences in what was valued and how different cultures and ethnic groups operate within practice contexts. I observed that it seemed important to use collaborative approaches to include different views and discuss with stakeholders what 'good looks like'. I also realised there appeared to be different dynamics operating within different practice contexts (such as power, devolution and regionalisation). I wanted to gain deeper understanding of such concepts and potentially contribute to this knowledge space and in practice.

I noticed that both managers and evaluators displayed variable understandings of results-based management concepts in different practice contexts. This led to me to reflect on what were the subject areas and epistemological concepts that informed practice and then to wonder whether there were possibly some common concepts that could assist practice. My management and evaluation studies provided some academic background to strategic planning and evaluation approaches in different contexts such as the architecture and use of results frameworks and use of theory-based evaluation.

## 1.2 Background

In the early 1990s, public sector reform internationally focused on changing the structures and systems of management with the development of new public management theories (Giddens, 1998; OECD, 2002). These reforms led to many countries focusing on improving their public sector economic performance rather than managerial reform (Ryan, 2003). These economic approaches sought to develop systems that would more efficiently and effectively deliver services for citizens at less cost to central government (Mayne, 2007a; Schick, 1996, 1998).

The concept of *Managing for Results* was first put forward by Drucker (1955), who emphasised focusing on the results (that is, changes) for customers, beneficiaries, and recipients, rather than the activities that were undertaken. In evaluation terms, this was the equivalent of a shift from outputs to outcomes (Perrin, 2006). Wholey (cited in Donaldson & Scriven, 2003, p. 45) considered that results-based management (RBM) emerged as a common element in the reform efforts to assist with demonstrating results of programmes and strategies, and to link activity performance to an organisation's goals and outcomes. White (2009) noted that RBM specifically focused on achieving outcomes rather than the inputs, activities, and output levels that characterised earlier approaches. These public management reforms were particularly evident in Canada, the United States of America (USA), the United Kingdom (UK), A/NZ, and Australia (Hughes, 2003, 2009) where extensive performance frameworks were introduced.

It was this results-based context that was the initial focus for my research. Practitioners and strategic planners realised that establishing and sustaining results-based monitoring and evaluation (M&E) systems needed an ongoing commitment (Kusek and Rist, 2004). These systems were viewed as providing "continuous reporting on findings can and should also extend to guiding decision-makers through implementation of recommendations" (p. 135). Initially, my research sought to put forward potential principles and concepts to support and sustain evaluative practice within the contexts of public sector management and the OECD (2005b). However, as my research process unfolded, so did my realisation that a results-based focus alone may not be adequate to support people working together in setting the strategic direction or management and governance of countries, sectors, and agencies to enhance more equitable and sustainable development.

Internationally from 2000 onwards, new paradigms in development, management, and aid emerged with the shift of focus to countries becoming the driving force of their own development and country systems (Segone, 2009). Multiple forums on aid effectiveness – many of which were coordinated by the OECD, resulted in the Monterrey

Consensus in 2002 (OECD, 2002) followed by the Paris Declaration on Aid Effectiveness in 2005 (OECD, 2005a). The Paris Declaration, initiated by governments particularly from the OECD and countries receiving development assistance, included five partnership commitments “to be interpreted in light of the specific situation of each partner country” p. 3) including “ownership, alignment, harmonisation, managing for results, and mutual accountability” (pp. 3-8). The latter, in particular required “managing and implementing aid in a way that focused on the desired results and uses information to improve decision-making” (OECD, 2005b, p.7).

Between 2006 and 2008, the OECD shifted the aid and development focus to working with country and sector systems. This shift in development thinking focused on national development and ownership (Conlin & Stirrat, 2008). Accountability demands were present, but there was also a shift from projects to broader programmes. This led to a widening of evaluation stakeholder interests aimed at shared objectives in country and sector development and management approaches. Results frameworks became a key component of the OECD Aid Effectiveness partnership commitments and *Managing for Development Results* agenda (OECD, 2007). The focus was on increasing aid effectiveness by enhancing the partner and donor countries’ efforts, through concentrating on strengthening partner countries’ national development approaches and frameworks” (OECD, 2009).

In the period since 2009, there were significant additional shifts made in country-led development, with donors aligning activities to government priorities represented in partner country frameworks. Historically, each development partner had their own prescriptive templates. However, there still appeared to be limited knowledge and description of how this alignment to country approaches could be achieved through planning, M&E activities given the nature of donor requirements (Corre, Mackie, & Trenner, 2008). By 2010, most development partners were using the term results frameworks (OECD, 2011). Yet there still appeared to be limited evidence of any systematic progress towards using frameworks to enhance development and aid effectiveness.

Increasingly, emphasis was being placed by countries on strengthening national-level M&E systems. Segone (2009) noted that if “evaluation systems were owned and led by concerned countries, it would facilitate the availability of more rigorous evidence such as country-specific data which is needed to monitor policy reforms and national development goals” (p. 23). He further stated:

It is fundamental that strong national evaluation systems exist to question national development policies. Without having strong national systems in place,

development results will not be sustainable, and evaluation will mainly serve external needs. (p. 23)

This research was undertaken, and thesis written from 2011 to 2020. During this period, Chimhowu et al., (2019) considered there was a relatively unrecognised global shift to 'new' national development planning in both developing and developed countries. A goal of this research was to contribute to the understanding of what makes for a strong national evaluation system alongside or integrated with national development, management and governance approaches.

### **1.3 Research context**

The real world (Bamberger, Rugh, & Mabry, 2006) focus for my study was the use of strategic results frameworks or the approaches through which a country articulates its organisation as an enterprise with operational and change activities directed towards national development strategies, and sector programmes. A key consideration that presented early in this research was the meaning of national development. I commenced my research cognisant of two different geopolitical contexts within which I wanted to undertake my research – namely developed countries (including A/NZ and Australia) and developing countries (such as PNG and Samoa). As I was undertaking the background reading (both academic and practice guidance), I realised there was a shift in focus towards more generic national development concepts rather than a distinction between an opposed view of developing countries (with a focus on donor aid and development) and developed countries (not reliant on international donor aid). Both contexts were adopting more convergent approaches.

This realisation assisted my framing for this research of country systems where results-based evidence sits in a national development, management, and evaluation nested hierarchy context. Visualising evidence gained through M&E activities as central rather than peripheral to development goals has been relatively explicit in donor-led development models since the 1980s. It more recently has come to characterise government approaches in developed countries under the auspices of evidence-based policy-making in the 1990s (Sutcliffe & Court, 2005). This thesis sought, through a theory building research lens (Lynham, 2002,) to develop a strategic evaluative approach that could be relevant in either context.

The focus of monitoring and evaluation (M&E) in a national development context is on collecting data about the progress towards development goals and policies generated by management activities and government functions, including interventions and other donor, non-governmental organisation (NGO), regional, and private sector activities.

Pawson (2006) presented a broad definition of evidence as “information gathered from a large variety of sources” which, once analysed, can be used to describe development progress (p. 15).

Evidence is an essential part of M&E, which in turn is a critical component of management, which in turn contributes to development, and there are reciprocal relationships between M&E, management, development, and governance (Betts & Wedgewood, 2011; Dahler-Larsen & Boodhoo, 2019). Given this convergence in an understanding of progress towards goals at a national level, I examined in more detail the impacts and significance of changes in the architecture and use of strategic results frameworks, evaluation, management, and governance in national development contexts.

## **1.4 Context of the study**

This research context focused on the impacts and significance of changes in the architecture and use of strategic results frameworks, the wider context of country development in which these frameworks are situated, and associated management and governance systems. Challenges in RBM and identified knowledge gaps were outlined in the literature review on the impacts and significance for development, management and evaluation knowledge and practice.

Both case study countries in my research were island nations based in the Pacific, have indigenous peoples, and were colonised, but they have different approaches to development planning. A/NZ has a devolved management approach to development strategy and evaluation, whereas PNG has had a more centralised top-down approach to development planning since independence in 1975. Both countries have devolved management accountability to agencies that received limited centralised guidance in evaluation. In addition, PNG receives international funding from development partners that adds to their accountability requirements. This research was undertaken in two phases: an initial scoping, followed by research on the two case studies of A/NZ and PNG.

Currently, countries and agencies rely on development models built around strategic results framework architecture (OECD, 2005b), as is the case in PNG, or around a central government architecture based on results-based public sector performance management, as is the case in A/NZ (Gill, 2008). To consider whether an alternative form of strategic accountability might be a feasible alternative to either of these approaches, this research focused on how to understand the impacts and significance of changes in the architecture and use of strategic or results-based planning and management and governance systems in these two contexts.

## 1.5 Conceptual research frame

My entry point into this research was to research how countries and sectors approached setting country and sector development goals, aligned management activities, used evaluation to measure an, assess and report their progress, and adapt to changing contexts and needs.

Strategic results frameworks with accompanying theories of change (Rolfe, 2019), are promoted as tools to aggregate results from multiple programmes, departments, and across agencies and sectors for making sense of country systems. These frameworks are significantly more complex than a programme or logic model (Rogers, 2008b). Results frameworks identifying intended results at both strategic and operational levels became recognised components of country-led M&E systems (Kusek & Rist, 2004; Rist, 2006). Strategic results frameworks and theories of change are tools that link strategic goals and outcomes to the lower programme or project goals and results which form the annual budget (Mayne, 2004).

This research began with a literature and guidance document review and scoping phase interviews to understand how results frameworks were being used as key tools in results-based management in the context of the guidance for their use issued by the OECD (2005b). The assumption here was that the ideas and interventions were represented as results frameworks at both strategic, sector, organisational, programme and project levels. How frameworks were developed was an initial focus of this research. During the review of literature and guidance documentation, I developed an impact model (Figure 2) based on Brinkerhoff's (2002) success case method which included concepts identified during my review of the interdisciplinary literature and guidance documents. The literature review discussion in Chapter 2 outlined these concepts included in the impact model based on the OECD (2005b) guidance within country systems. During my literature review, I noted there were gaps in the knowledge space and areas requiring further research observed in the literature. These gaps and areas for further research were highlighted in Chapter 2 of this thesis.

This research was underpinned by critical realism philosophy as outlined by Bhaskar (1979/1998) and Wuisman (2005) (see Figure 2). I selected critical realism as an underpinning theoretical perspective to examine different ontologies in country systems. The critical realist framework enabled me to critically examine and analyse the epistemology and emergent good practice concepts from different perspectives (managers and evaluators at central agency and line agencies that may be operating with country and sector system) in different country contexts. I consider that when I began my research, I was relatively uncritical in my thinking and tended to simply accept

recognised practice concepts. However, as I built my theory, my thinking was robustly challenged, and I also came to see opportunities to further evolve the critical realist framework. During the scoping phase, I found there were additional good practice concepts that I had not included in the initial impact model, but which interviewees considered were important: information technology, systematic inclusion of stakeholders in collaborative approaches, and increased managers' evaluative capacity.

After closer examination, I considered that these additional concepts could be grouped with aspects of Pawson's (2006) contextual layers. For example, "information technology" aligned with Pawson's conception of a layer of (i) "ideas, interventions or frameworks" and "systematic inclusion" was relevant to his concept of (ii) "within institutions such as public sector agencies", and "increased managers' evaluative capacity" could be inserted into his (iii) "wider infrastructural context". This allowed me to adapt Pawson's existing model to accommodate ideas that were emerging from my research, and develop a new typology with three contextual layers for my theory-building research in national planning, management and evaluation.

My sense was that my interviewees considered interpersonal relations and individual capacity (another two of Pawson's contextual layers) were operating differently within and between the three layers. This led to me reflecting more critically on what was in the literature and where potential knowledge concepts could emerge from this research. I decided to use and adapt aspects of Pawson's contextual layers as an emergent 'realism' research conceptual frame for my research. I liked that the three contextual layers were a starting place to structure my qualitative findings from my research, but it was not a perfect fit. Up to this point, I consider I was rather uncritical academically and as a practitioner regarding literature and guidance documents. Yet, I was aware that there was potential as part of this research to examine national planning, management and evaluation knowledge and practice concepts more closely and considered this could be a knowledge space where this research could contribute.

Vasilachis de Gialdino (2011) refers to the qualitative research journey as shifting from "knowing the subject" to "known subject" (p. 2). I found this research journey challenging as I researched managers' and evaluators' views and artefacts about what I considered was an object with me 'knowing the subject'. I tried to be objective as a researcher. However, my emerging findings challenged what I thought was a known reality. I then had to accept, as a researcher and practitioner that I needed to acknowledge my research reflections in my theory building I approach and analysis during this research journey. I struggled to make sense of different countries' approaches to national planning and what a realist ontology was on the ground. Vasilachis de



Gialdino refers to using reflection points during the research journey as “epistemological reflections” (p. 3) where I considered academically and as a practitioner about my research progress. I realised how human agency and contexts seemed to combine to form different ontologies as PNG and A/NZ approached country systems and national planning in almost diametrically opposed ways. PNG uses centralised planning and coordination with regional inclusion mandated. Whereas A/NZ devolves country and sector planning and management to line agencies with limited centralisation and coordination, and had no mandated inclusion of Māori until late 2020, apart from Treaty of Waitangi principles.

I questioned the legitimacy of me as a researcher and practitioner inserting my academic and practitioner ‘lived’ experiences into my research. However, I felt I could adapt Pawson’s contextual layers (based on my scoping phase findings) and develop my own epistemological typology. Therefore, before undertaking my case study research, I developed this typology during the scoping phase and used this my own framework throughout the case studies phase, and the analyses.

As my thinking developed, I realised I could use the critical realism strata of experience, events and mechanisms (Wuisman, 2005) as a secondary analysis frame. This assisted me to examine how the concepts appeared to operate within and between the three contextual layers. My aim was always to identify potential good practice concepts and their potential contribution within country systems to enhance development effectiveness.

As I progressed along this theory-building research journey, I used my practitioner and research experience to iteratively shift from ‘knowing the subject’ to what became my ‘known subject’. This shift entailed developing a unique typology through which to assess country systems across three contextual layers (as identified by Pawson and others) but to also highlight the contribution of embedded human agency and context concepts as central. Thus, while the research embarked on a critical realist it diverged significantly from Pawson’s original conception. While Pawson was an important sounding board, this thesis could not be considered as part of the purist ‘critical realist’ approach.

## **1.6 Research questions**

My overarching research question was: *What good practice evaluative principles and concepts may enhance country, sector and agency development, management and governance knowledge and practice in different contexts?* From the conceptual framing outlined above, the following sub-research questions were refined as part of an emergent research design:

1. What evidence from A/NZ and PNG demonstrates what works for whom in relation to results frameworks and associated management and governance systems? (RQ1) (Chapters 5, 6 and 7)
2. How and in what ways can conceptual links be identified between frameworks, management and governance systems and good practice principles? (RQ2) (Chapter 8)
3. How and in what ways can country, sector, and agency strategic results frameworks and their associated management and governance systems underpin the development of good practice values and principles? (RQ3) (Chapters 8 and 9)
4. How do the emerging good practice principles in country, sector, agency and programme systems work to enhance development effectiveness and governance for development (national governments and partners), management, and evaluation? (RQ4) (Chapters 8, 9, and 10)

These questions guided research undertaken from 2011 to 2016, across two comparative case studies (one at a deeper level in A/NZ and the other at a lighter level in PNG). These questions provided the basis for the analysis of the findings and the identification of good practice concepts and principles. The implications for management and evaluation knowledge were then considered.

## **1.7 Problem statement**

The use of country-led M&E systems, incorporating results (Kusek & Rist, 2004) or outcomes frameworks (Duignan, 2008) aimed at enhancing evidence-based policy and decision-making in national development contexts. However, Segone (2009) considered that the knowledge base for country-led and owned M&E systems was “only slowly growing” (p. 4). Corre et al. (2008) also questioned the level of partner participation in evaluation. Ba Tall (2009) likewise advocated identifying, creating, and using knowledge while building capacity.

Vedung (1997) observed that moving to a result or outcomes focus in the public sector confronted managers with new problems. These included the monitoring of programme outcomes, and the identification of the role of evaluation in results-based management, which required more sophisticated approaches than previously used. Several critical issues noted as impacting evaluation demand included organisational thinking, learning capacity, and institutional weaknesses.

Ryan (2003) highlighted potential issues for A/NZ based on the 1988 Commonwealth of Australia's *Evaluation Strategy* in taking a ‘big bang’ approach to evaluation. These

issues included the long lead times to set up and establish agency-wide M&E systems. He considered that in the short term, more could be gained in A/NZ on “creating evaluative management cultures inside organisations” (p. 6) emphasising learning, and strategic internal evaluation focusing on outcomes. Ryan, in 2011, further emphasised that little evaluation appears to be undertaken by government departments in A/NZ, which aligned with Segone’s (2009) observations about the slow growth of the role of M&E.

Stern (2008a) considered that an increased focus on country and sector systems, underpinned new paradigms emerging in evaluation, was needed highlighting the principles of “ownership and accountability” (p. 3). Increasingly, governments, managers and communities (Buffardi, 2016) needed to be able to adapt more responsively for resilience and sustainability, to enhance people lives. A programme example aimed at building resilient development in the Pacific was a suite of projects financed under the *Pacific Risk Resilience Programme* (UNDP, 2016). This programme used integrated and adaptive approaches focused on partnerships and participation to enhance development outcomes.

Some of the issues I identified from the research scoping discussions and initial literature review based on my academic studies and practitioner experience included agencies using multiple unaligned frameworks, operational silos within public management line agencies, and limited understanding and knowledge sharing by managers within and between functions and roles. Evidence of the impact of these issues could be identified in practice and pointed to knowledge gaps within the development, management and evaluation approaches used by managers and evaluators within different development settings – country, sector, region, programme and communities. Increasingly, I became interested in trying to think through what kinds of knowledge were missing in management and governance contexts that allowed these kinds of perverse processes to persist and whether a different strategic-level approach might make a difference.

From my own experience and practice, I wondered whether different knowledge and practice concepts might more usefully underpin the interdisciplinary interface between development, management and evaluation knowledge fields than what was currently recognised. This led me to examine approaches, values, principles and good practice concepts – in particular, those that related to issues of inclusion, development effectiveness, governance and sustainability for communities, organisations, regions, sectors and countries.

Hence, my initial thinking behind this research was that there needed to be a better way to connect evaluative approaches within management and, potentially, with the new SDGs in international and national development. The research direction evolved more explicitly towards an aim to contribute to the management knowledge base through theory building by identifying good practice principles (conceptual and operational). This was to extend knowledge concepts and support an expanded role for evaluative perspectives and practices that could be more effectively used to enhance inclusive and participatory forms of development governance within national planning contexts.

## 1.8 Research timeline and thesis structure

The theoretical framing, research phases, emergent theory-building methodology and analytical frames for this research are outlined in the timeline and thesis structure diagram (Figure 1) and thesis structure description.

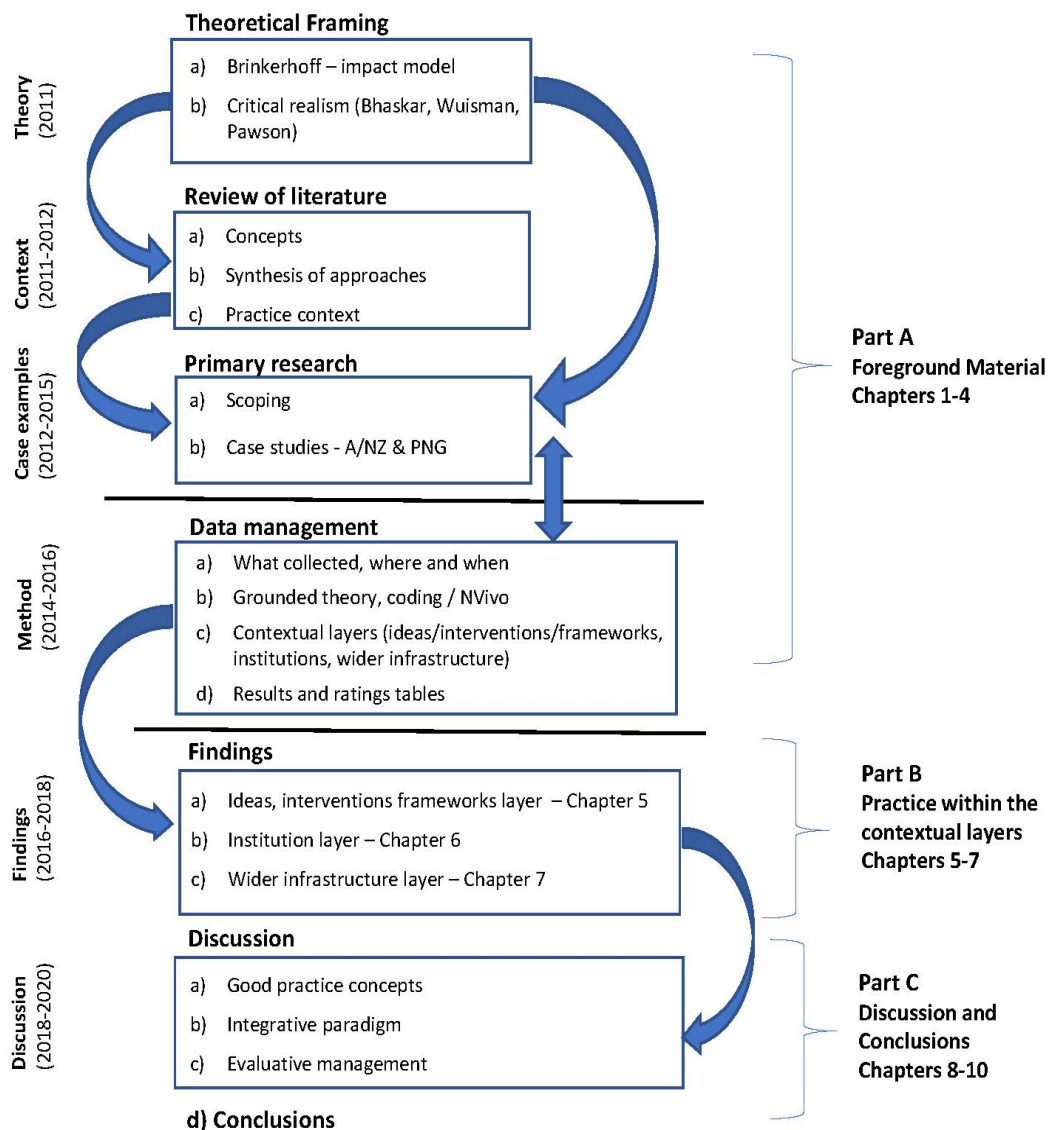


Figure 1: Research timeline and thesis structure.

This thesis was organised under three parts with the following chapters.

#### **Part A: Foreground material:**

- **Chapter 1: Introduction** – includes starting place for research, the background and context to the research, problem definition, and research questions and phases.
- **Chapter 2: Research knowledge context** – covers key concepts in the development, management and evaluation knowledge fields, including thematic knowledge fields and concepts, and synthesis of knowledge gaps.
- **Chapter 3: Methodology** – outlines how the research was undertaken in the research scoping phase and two illustrative case studies A/NZ and PNG.
- **Chapter 4: Practice contexts** – includes discussions of A/NZ and PNG country contexts.

#### **Part B: Findings – Practice within the contextual layers:**

- **Chapter 5: Ideas, interventions and frameworks layer** – substantive and emergent findings, including individual capacities of managers and evaluation practitioners.
- **Chapter 6: Institution layer** – substantive and emergent findings, including interpersonal relations, technology, performance management and reporting.
- **Chapter 7: Wider infrastructure layer** – substantive and emergent findings, including nature and role of development (national governments and agencies), management, evaluation and governance.

#### **Part C: Discussion and conclusions:**

- **Chapter 8: Concepts linking frameworks, management systems and principles** – critical discussion of the substantive and emergent findings, development of an idea for repositioning strategic evaluation, repositioned identifiable good practice concepts, values and high-level principles supported by additional literature.
- **Chapter 9: Emergent new integrated management and evaluation paradigm**, referred to as *evaluative management*, and its application to other settings is discussed. An integrated *Values-Based Evaluative Management System* model is put forward for consideration, underpinned by values and high-level principles.

**Chapter 10:** Conclusions and implications for *Values-Based Evaluative Management* and associated good practice concepts are outlined. Consideration of the implications for development, management, and evaluation knowledge is undertaken, and areas for further research are identified.

## Chapter 2: Research knowledge context

In this chapter, the key literature that provided the context for my research is rehearsed in six sections. First, some of the contested terminology is discussed. In the second section, the concept of development provides the focus. Publications from the international agencies, such as the United Nations and the World Bank are central to the discussion. The third section considers what was written about development by evaluators and commentators, with a focus on the practice of evaluation. Section four examines management literature to provide background literature that outlines predominant management theories and approaches. Section five displays the research impact model, — the initial conceptual frame for this research that was used during the scoping phase. (The emergent research conceptual frame used for the cases studies is discussed in Chapter Three). The final section of this chapter brings the development, evaluation, and management perspectives together to identify existing themes and gaps in the literature and debates.

My research was situated at the multidisciplinary interface between development, management, evaluation and governance for national planning, management and evaluation. This research was framed within a concept of a “country system” where focus on accountability for development activities, and goals were increasingly being placed on countries to plan and evaluate their development results (Kusek & Rist, 2004; Segone, 2009; Picciotto, 2009; Betts & Wedgwood, 2011; Schwandt, Ofir, D’Errico, El-Saddik, & Lucks, 2016; Geoghegan, et al., 2019; Chimhowu et al., 2019). In each of the knowledge fields – development, management and evaluation – key theoretical approaches and practice concepts were sought that reflected ideas of ‘good practice’ or recognised knowledge gaps and what I needed to do was identify these concepts from across the broad literature and place them into an analytical framework that could provide an overview.

There were two key challenges encountered during the initial review of literature and the iterative updating of literature during the period of the research. The first challenge was dealing with the breadth of literature that interfaced with my research topic. I worked to overcome this by scanning a breadth of literature and noting key knowledge concepts and then by keeping up with recent literature in these areas, whilst scanning for related literature on emergent findings from my research.

Secondly, from 2012-2018, during my epistemological reflections on theory-building, I often struggled with the ontology of the research object and topic. I consider this was due to my uncritical thinking based on some practitioner experience at the initial phase

of this research where I thought the research object was more definitive than it appeared to be in the literature and in research. Unbeknown to me, a group of academics working on a collaborative project were also interested in the trend I was seeing toward a re-emergent national planning approach. When Chimhowu et al. (2019) published their initial research into national planning, I felt relieved. They outlined how national planning was a relatively unrecognised and noted the re-emerging trend towards it along with some countries using more collaborative approaches. I could see this trend evident in practice in countries such as Australia, Canada, Indonesia and Papua New Guinea, as I persevered with my theory-building research.

From 2018 onwards, with an increased interest globally in country sustainable development, my research appeared to be converging in part with these interests. However, the prevailing management paradigm — an economic focus with some tagged on social and cultural considerations — appeared to remain somewhat intact. The social/cultural and environment dimensions of sustainable development seemed slow to significantly influence mainstream management theory and practice. My belief is that the impacts of climate change, Covid-19 and other potential pandemics, and growing inequalities may, in time, cause or force a renewed focus on human agency and country context for national planning, management and evaluation systems. In the meantime, I contend that a model, such as the one I outline, provides a possible approach to including such perspectives in development planning.

## **2.1 Approach to the literature analysis**

Given the breath of this interdisciplinary research topic, I designed an organic process using a matrix table to assist my identification and synthesis of key concepts. Then I developed an initial literature concept map (Appendix A) as concept topics emerged. First, I put the starting topic for this research (Architecture of Results and Outcomes Frameworks) at the centre of the map. I placed the three knowledge fields (Evaluation, Development and Governance). As concepts emerged, I added these to the map situated within the relevant knowledge fields. Given the breadth of topics identified in the management field, I decided to separate management concepts from governance within my literature review. The development of the concept map provided structure for grouping knowledge topics identified during the literature review, particularly given the interdisciplinary nature of this research. In addition, the mapping assisted me to expand my awareness of potential concepts that might emerge from the case study fieldwork. I used a template to document key points and then undertook a thematic analysis to identify key gaps. A synthesis of the knowledge gaps was undertaken, and my thematic

analysis also showed knowledge themes requiring further research. I then iteratively included these themes in my literature review.

In my initial broad-brush reading, I looked for work on strategic frameworks and their historical development within evaluation (Chen, 1990; Rogers, 2000, 2008; Weiss, 1972), management (Aucoin & Jarvis, 2004; Hatry & Lampkin, 2003; Hughes, 2009; Perrin, 2002) and development (Gasper, 1999, 2001). I followed back through the history of framework development (Binnendijk, 2000, 2001), national development (Chambers, 2010; Chimhowu et al., 2019; Ferguson, 1990; 2002; Hulme, Savoia & Sen, 2015; Meier, 2005; Pieterse, 2010; Rodril, 2011; Scott, Joubert & Anyogu, 2005), the history of the OECD (OECD 2005a&b, 2006a&b, 2008a,b&c, 2009a,b,c,d,e,f,g,h&i, 2010a,b,c&d) and development partners' use of frameworks within management (Knack, 2013; Leonard & Bayley, 2008; Moynihan, 2005, 2006). I looked at a range of literature on the role and nature of evaluation within different development settings (Bamberger, Rugh, & Mabry, 2006; Beloe, 2005; Conlin & Stirrat, 2008; Dabelstein & Patton, 2013; Davies, 2004; Hummelbrunner & Jones, 2013; Kindornay & Morton, 2009; Mayne, 2007a; Mackay, 2007; Picciotto, 2002, 2009; Vedung, 1997). However, while this suite of material provided useful insights on the architecture of frameworks with development settings, there appeared to be knowledge gaps on the application and use of strategic results frameworks with stakeholders within management and governance contexts.

The research then included an iterative focus on management and evaluation literature such as Mintzberg (1985, 1994), Grant (1998), and Kaplan and Norton (1996); results-based management in particular (Kusek & Rist, 2006; Mayne, 2007b; OECD, 2017, 2019; Poate, 1997; White & Rodriguez-Garcia, 2009; Wholey, 1999,); performance management (Aucoin 2012; Aucoin & Jarvis, 2004; Newcomer, 1997); theory-based evaluation (Chen, 1990; Duignan, 2007; Funnell & Rogers, 2011; Rogers, 2000, 2008a&b); goal-free evaluation (Scriven, 1991, 1997); realist evaluation (Pawson & Tilley, 1997, 2001); empowerment evaluation (Fetterman & Wandersman, 2005; Stufflebeam, 2001); responsive evaluation (House, 2001; Stake, 2004); internal evaluation (Mathieson, 2005); strategy evaluation (Patrzi, 2010); Mātauranga Māori and global indigenous knowledge (Macfarlane, 2016); Kaupapa Māori evaluation (Carlson, Moewaka Barnes, & McCreanor, 2017; Cram, 2016) and systems and complexity concepts in evaluation (Larson, 2018; Renger, 2015; Williams, 2016; Williams & Iman, 2007).

Two particular elements that stood out from this second wide tranche of reading were the use of systems thinking in evaluation approaches and in indigenous evaluation. While Pawson's (2006) critical realist framework provided an initial basis for my thinking about



the contexts for RBM, the motivation to develop my own typology for this research of country systems derived, in part, from the insufficiencies in existing theorising about indigenous or non-western approaches. The concepts and ideas I encountered through reading about indigenous evaluation noted an emphasis on incorporating values, particularly collaboration, inclusion and participation throughout planning and implementation of services and interventions to enhance inclusion and equity. This led to me giving more explicit consideration to the idea that different stakeholders would have different ontological views and values and, in turn, that the failure to acknowledge such views could give rise to biased and inadequate approaches.

Further iterative reading called me to pay more attention to emergent themes of organisational capability (Arygris, 1997; Canadian International Development Agency [CIDA], 2007; Grindle, 2007), capacity building (Grindle & Hilderbrand, 1995; Sen, 1999, 2002), and governance (Betts & Wedgwood, 2011; Kaufman & Kraay, 2002; Kaufmann, Kraay, & Mastruzzi, 2010; Peters, 1998). Finally, the review was expanded and updated to include an examination of management and evaluation approaches in central and government agencies (Bamberger, 2000, 2010; Dorbell, 2002; Picciotto, 2002, 2009 Ryan, 2003; Schick, 1996, 1998; Stern, 2003, 2008a&b) and broader infrastructure practice contexts (Peters & Savoie, 2000; Chakravartty, 2007; Dahler-Larsen, 2012, 2019; Feinstein, 2017; Geoghegan et al., 2019; Schwandt et al., 2016; Rai, Smith, & Brooks, 2019; Thomas, 2002).

My analysis noted there were convergent good practice themes on management and governance, which included a use of frameworks, strategic, performance management indicators and data, and reporting. However, the evaluation approaches showed a history of earlier divergence on evaluation approaches that predominantly focused at operational and programme levels with a noted gap on strategic level evaluative approaches where systems approaches were increasingly incorporating aspects of complexity sciences (Mowles, 2014) to enhance development effectiveness and governance.

Based on this expansive review of development, management and evaluation literature and documenting key concepts in a matrix, I undertook a thematic analysis. There were key concepts that emerged which as an experienced practitioner, I considered were relevant given the interdisciplinary interface and nature of this research. What was surprising to me was the commonality of themes that related to the different levels such as the architecture and use frameworks, institution and wider infrastructure. Yet from my literature review, I still struggled to see how evaluation interfaced with performance management at a strategic level within institutions and the wider

infrastructure contexts for sectors and countries. In addition, I could see there were differences in the use of terminology between development, management and evaluation literature which appeared to be adding to the complexity.

In each of the following sections, I highlight earlier and more recent literature that helped inform my thinking, analysis and writing of this thesis. Some of the more recent literature listed in this section is related to the emerging findings and was added iteratively during the writing up of my thesis. This was part of my theory-building approach. Relevant literature is also referenced directly in my discussion sections in Chapters Eight and Nine.

## **2.2 Terminology**

At the beginning of my research, I looked at the specific terminology and concepts that have become part of the discourse of national governments and international development agencies such as the OECD, under the auspices of the Organisation for Economic Co-operation and Development – Development Assistance Committee OECD-DAC, in order to work within existing definitions. These terms include concepts such as “country systems”, “sectors”, types of “frameworks”, and “governance”. While there is consistency in terminology across international development agencies, there is no formally agreed terminology across development, evaluation, and management knowledge fields. Some terms are universally accepted such as RBM (OECD, 2010b) and other terms such as “frameworks” (country, sector, agency, results and performance) have different meanings to different stakeholders and users. This causes confusion in the literature and practice and requires some pragmatic response.

Given changing paradigms in development, evaluation, and management with an increasing emphasis on SDGs (UN, 2015), some terms now appear to be redefined to have a more common understanding. My discussion of the literature built on the terms included in the glossary developed by the OECD (2010b) and included definitions and discussion for terms which were used in this research that appeared to be more contested. In particular, the descriptive discussion that follows looks at development; aid and development effectiveness; governance, accountability, and Results-based Management. The summaries, the first of which is “Development”, provided a narrative both about key concepts but also about broader political and social changes that a range of commentators have thought to be relevant to a public understanding of these key terms.

## 2.3 Development

Development is regarded as a “complex struggle” (Pieterse, 2010, p.xviii). Watts (2006) outlined that the concept of developmentalism encompasses a “wide-ranging and philosophical position underwritten by panoply of development institutions, discourses and practices – that sustains the idea of development as a normative goal” (p. 123). This intended transformative process was underpinned by a desire to improve parts of the world referred to as the developing world also referred to as the “South” (UN, 1987, p. 49) or the “Third World” (Rist, 1997). Underpinning this development process was an “idea that developmentalism is best pursued through an interventionist, centralised, bureaucratic, and authoritarian state that self-consciously governs the market and the process of capital accumulation” (Watts, 2006, p 123).

However, in the 1980s, a shift to neoliberalism and free-market economy appeared to challenge this post-1945 view of development. For example, Rostow’s 1960s model (Rostow, 1960) outlined stages of growth for economic development. This increased focus on economic development under neoliberal economics had underlying assumptions that the market would provide adequately for citizens. However, questions were raised about the value from the “development apparatus” (for example in Lesotho by Ferguson, 1990, p. 7) where development organisations such as World Bank “generate their own form of discourse” (p. xiv) and Ferguson queried what is real development?

Pieterse (2000) emphasised that it is important to be aware of whose discourse it is and how real it is. A key focus for my research was examining how different stakeholder perspectives (particularly indigenous) were considered in post-colonial states (Gupta & Sharma, 2006) which included countries that were colonised based on western hegemony (Pieterse, 2000) within indigenous contexts. Therefore, gaining an understanding of discursive shifts in development planning, management, governance and evaluation over time in different contexts was important for me to keep the ideas being captured in my matrix in perspective.

As development thinking changed over time, the idea of progress, as put forward in classical political economy in the 1900s (UNDP, 2015a&b), also changed. The idea of human-centred development with the Millennium Development Goals (MDGs) evolved at the turn of this century with the inclusion of sustainable development as one of the 10 goals. The MDGs identified that:

achieving poverty eradication and sustainable development will require significant transformations at all levels. These will include the strengthening of policy integration and coherence, adaptation by institutions, fostering of effective

multi-stakeholder partnerships and enhancing follow up and review of implementation. (UNDP, 2015a, Agenda item 5 (c) p. x)

A shift towards social development placed more emphasis on longer-term policy-making (Haggard & Kaufmann, 1994), which impacted on development practitioners realising that achieving development results required longer time-frames. Brinkerhoff (1996) emphasised that policy makers and public managers faced the challenge of “sustaining these reforms beyond the launch phase so that policy changes, whose benefits rarely appear in the short term, can bear fruit” (p. 139). However, Pieterse (2010) observed that “the strength and weakness of development thinking is that of its policy-orientated character” (p. 12). He argued that development is problem-driven as opposed to being theory-driven which can result in power imbalances between communities and governments.

In 1992, the United Nations and the OECD agreed to a focus on National Sustainable Development for developing countries as part of *Agenda 21*, and the major agreement was signed at the Rio Earth Summit (OECD, 2006) This agreement included social, economic and environmental strategies to ensure “socially responsible economic development for benefit of future generations” (p. 145). There were also challenges noted around human-centred progress with the MDGs for example, by Barrett (2011) in education (where constraints were highlighted in accessing education).

Pieterse (2010) surmised that “long-term trends in development, parallel general shifts in social science” (p. 12). He considered that from the 1990s economic drivers underpinned country social development approaches and policy. Chambers (2010) concurred that by 2000 the participatory approaches of the 1990s had shifted to focus more on accountability.

During 2012–2016, there was a shift in the global development context from Millennium Development Goals (Riechert, 2006; Sachs, 2012) to SDGs (UN, 2015). The significant shift to SDGs led by United Nations (UNDP, 2015a&b) was a response to an increasing emphasis on global and local level environmental, social, and economic concerns. Chimhowu et al., (2019) in their research on national development planning observed that the production of national development plans had doubled from 2008 to 2018 and varying approaches included “technocratic, collaborative and social embeddedness” (p. 84) which they noted were emerging concepts within this ‘new’ national planning paradigm. Increasingly, countries (government and citizens) in both developed and developing contexts want and need to respond more proactively to social, economic and environmental resiliency concerns and improve their development results and lives for their citizens.

Questions were becoming stronger by countries and their citizens over whose worldview (Vidal, 2008) and whose development is it? A shifting development focus at country level appears to be supporting responsiveness to indigenous perspectives in polities such as PNG, Samoa, A/NZ and Australia. This approach includes expanding national indigenous capacity and capability to manage their own development (Ba Tall, 2009) and address the impacts of colonisation (Crocombe, 2001; Cram, 2016). This growing awareness of country, and indigenous stakeholder perspectives within international and national development, management and evaluation contexts were relevant to my research as I focused on development at a country level and the strategic role of evaluation within its public management, evaluation and governance approaches.

As my review of literature progressed, it highlighted for me that development partners and countries appeared to have different perspectives on development and that countries such as PNG and Samoa value their own views on their development. Moreover, indigenous development thinking was emerging as a separate view of development. This led me to question how does A/NZ view its own development with its multiple Māori, Pacific, European, Asian population groups? And how are the different perspectives considered and included in planning, management and evaluation?

## **2.4 Management and accountability**

In public administration, the 1990s were a decade of major changes with a strong drive for enhanced transparency and increased accountability, which aimed to strengthen institutions and enable more effective government. Traditionally, accountability of public funds was regarded as the role of government agencies to provide to central government with underlying assumptions the services and interventions are reaching the beneficiaries – public citizens. The rationale for these changes was that public administration had “become isolated and out of touch” (Peters & Pierre, 1998, p. 4), from economic pressures.

The role of strategic planning was supported by Mintzberg and Waters (1985) to address gaps in strategic thinking and strategy development. They emphasised the value of mapping the intended results from strategy and updating “strategic intent” during implementation from “deliberate to emergent strategy” (1985, p. 257). However, Mintzberg (1994) was critical of formalised strategic planning processes as he considered that it can impede strategic thinking. Power (2011) outlined that Mintzberg considered:

analytical planning processes can deceive managers into thinking that they are planning strategically, and hence improving future organisational performance.  
(p. 1)

Quensel (2009) supported Mintzberg's view of strategic planning as he reinforced that "strategic intent is foundational to country-led M&E systems" (p. 6). Quensel outlined that strategy was developed to achieve a goal or provide a solution, and it is the management of activities that were required to make progress. M&E can then provide the data and evidence to enable the assessment of the results. Pal and Teplova (2003) considered that most OECD countries had performance measurement and planning regimes. However, internationally, governments were facing calls for accountability, and managers were asked to demonstrate development results to their internal and external stakeholders (Aucoin & Heintzmann, 2000). The contemporary reform effort to improve accountability was also part of a trend towards using results-based management to improve measurement of results and outcomes (Aucoin & Jarvis, 2004).

Moreover, Dwivedi, Khator, and Nef (2007) considered that in post-colonial settings, public sector management processes were impacted upon by legacies of bureaucratic administration processes from colonisation. These processes hinder countries and sectors, adaptive planning responses and contribute to accountability and transparency issues. These issues were relevant to this research as their colonial legacies influenced both PNG and A/NZ's management processes. The practice contexts for the two case studies in this research are outlined in Chapter 3.

Public sector reform also raised accountability issues about the delivery of public services as agencies wrestled with the challenges of shaping and managing viable and responsive public sectors (Schick, 1996; 1998). Robinson (1999) argued that the task of public management was one of managing an area of public action that included a range of actors and agendas. Also, the drive for increased accountability in public sectors appeared to be linked to private sector management, where there were specific standards and measures of performance (Aucoin & Heintzmann, 2000). From 2000, the public wanted to see this information included in decision-making and in response to their needs and concerns (Aucoin & Jarvis, 2004). Kaufmann et al., (2010) emphasised the importance of the links between accountability and development results and advocated for increased governance.

## **2.5 Results-based management (RBM)**

According to Wholey (cited in Donaldson & Scriven, 2003), RBM emerged as a common element in current reform efforts and was designed to assist with demonstrating results of programmes and strategies, and linking activity performance to organisations' goals and outcomes. White (2009) noted that RBM specifically focused on achieving outcomes rather than on the inputs, activities and output levels of earlier approaches. There was published consensus on aspects of the key approaches and steps for results-

based management (Duignan, 2009; Kusek & Rist, 2004; Pawson, 2006; Rogers, Hasci, Petrosino, & Huebner, 2000).

New Zealand's Treasury and the State Services Commission (2005b) emphasised that national audit agencies in Australia, Canada, A/NZ, the UK and the USA concurred on the seven board aspects for agencies to successfully manage for outcomes (results), which included:

- have a clear vision of why they exist, what they want to achieve and how well they are achieving against this
- plan their work keeping in mind a clear set of objectives, activities, outputs, outcomes and measures of success
- deliver what they have planned, and in a manner that is consistent with public service ethics, values and standards while meeting standards of timeliness, quality, accuracy, etc.
- take stock of their progress by monitoring, measuring, reviewing and evaluating as they go
- learn from success and failure and modify what they do and how they do it in response
- report publicly on their results, promoting transparency to Parliament and the public and providing a basis for dialogue about future decisions
- have an adaptive and innovative culture and seek continuous improvement. (p.1)

Strategic and programme results or outcomes frameworks were increasingly being used to provide the architecture or structure for evidence to inform decision-making. These approaches built on the Logical Framework Approaches (LFA), which originated in the 1970s (Gasper, 1999, 2001) from aid programmes. Other key themes on success for RBM that emerged from the international literature and practice guidance included the importance of leadership, time required for success, a need to create demand for results-based M&E information, and the need to align with existing budget and other systems (Perrin, 2002a; Leonard & Bayley, 2008). Other necessary preconditions Leonard and Bayley noted for success included "a demonstrated commitment in management decisions, an incentive structure that encourages public servants to operate in the public interest, and a degree of stability in staffing" (p. 18).

Overall, from my review of results-based management literature, I found there were recognised aspects of results-based management that could be applied to both

developed and developing countries. However, the availability of financial resources was a noted constraint for developing countries (OECD, 2019).

### **2.5.1. Aid effectiveness and architecture**

In the 1990s, the primary focus was on aid effectiveness which traced back to the 1995 OECD-DAC and the Monterrey Consensus. These forums focused on supporting the growth of developing countries and an increase in living standards. Since 2002, multiple conferences and initiatives, including the First, Second and Third International Roundtables on Managing for Development Results (MfDR) and included progressive roundtables in Washington, DC 2002, Marrakech 2004, and Hanoi 2007, and the formation of the Joint Venture on MfDR (OECD, 2005b) moved this agenda forward. In Marrakesh (OECD, 2006a), the development agencies endorsed five core principles for MFDR, which included:

- i. focusing the dialogue on results at all phases of the development process
- ii. aligning programming, monitoring and evaluation with results
- iii. keeping measurement and reporting simple
- iv. managing for, not by, results
- v. using results information for learning and decision making. (OECD, 2006a, p.3)

However, Christiansen and Rogerson (2005) considered:

The pattern of aid actors, spending and activities we are presented with today are the product of disjointed history of political initiatives based on widely different premises. The current aid system is not a coherent, functional structure based on a single design or even on commonly held principles. (p. 1)

Despite warnings such as this, the OECD Paris Declaration of Aid Effectiveness principles with an increasing emphasis on development effectiveness became foci in development from 2005 onwards in both developing and developed countries. The OECD (2005a) principles aimed at providing the structures (that is, the architecture) and mechanisms for development coherence which Corre et al., (2008) considered were needed to align country priorities with multiple development partners. This shifting of focus from aid donors to country-led development was relevant to my research as it raised awareness and increased support for countries such as PNG in their development.



### **2.5.2. Development effectiveness**

Increasingly, aid architecture reoriented more towards countries' development priorities alongside funder priorities. The focus was on increasing development effectiveness by enhancing the partner and donor countries' efforts, which concentrated on strengthening partner countries national development strategies and operational frameworks. Kindornay and Morton (2009) outlined that development effectiveness included "organisational effectiveness, coherence or coordination, development outcomes from aid, and overall development outcomes" (p. 1). As part of governance and development effectiveness, results frameworks were promoted as a critical component of the OECD-DAC aid effectiveness partnership commitments and managing for development results agenda. Conlin and Stirrat (2008) noted a further shift in development thinking towards an increased focus on national development and ownership which included "shared objectives" (p. 201) with stakeholders.

These changes of emphasis led to challenges for public officials in national governments and development partnerships on how to link increased accountability, measuring of results, and transparency to decision-making. These trends were relevant as my case study research in PNG examined approaches and relationships between the PNG Government and its development partners.

### **2.5.3. Results frameworks**

By 2000, results frameworks had become recognised components of country-led M&E systems at both strategic and operational levels (Kusek & Rist, 2004; Rist, 2006). Increased emphasis on the evidence-base of results and outcomes from sector and programme activities. M&E activities based on results frameworks under the Managing for Results agenda were tools for both public sector and aid agencies to assess performance. (Australian Government, 2011). Binnendijk (2000) described the use of results frameworks as:

Results frameworks are useful as strategic planning and management tools. They can help a donor agency's operating units at the country level to clarify the significant development objectives they seek to contribute towards, in collaboration with partners, and to develop a logical plan or program strategy for their part in achieving them. Over time, operating units can begin to align (focus and concentrate) their activities within coherent programs that support the specific intermediate outcomes for development. (p. 35)

Binnendijk (2000) promoted the use of results frameworks at country and sector levels and for programmes in both developed and developing contexts. These approaches were further refined by Kusek and Rist (2004) in their ten-step approach to designing

country and sector M&E systems. An issue in terminology here is that terms results and outcomes were used interchangeably in some settings. For example, Duignan in A/NZ (2004, 2010) influenced outcomes thinking in A/NZ and, in Australia, Funnell and Rogers (2011) applied concepts such as theories of change and logic models that were becoming more widely used internationally in the UK, Canada, USA and A/NZ. as “an integral part of the evaluation process” (Cram, Hopson, Powell, Williams, & Kaul, 2019, p. 2)

From 2005, funders increasingly called for greater robustness in the evidence used to measure results (or outputs and outcomes). As a response to this international drive for evidence-based policy and decision-making (Pawson, 2006), some developed and developing countries changed their approaches to planning, policy, monitoring and evaluation. Donors moved towards a system of monitoring and reporting outcomes, to better plan towards, and assess the impact of aid efforts (Kotvojs, 2006). Leonard & Bayley (2008) outlined that strategic and programme results frameworks in different contexts included three key components: the results framework, measurement table and plan.

These RBM approaches were endorsed by the former Australian Government Agency (AusAID) and the New Zealand Agency for International Development (NZAID) from 2010 with guidance documentation produced. From 2012 onwards, these two agencies merged retrospectively with their Trade and Enterprise agencies and their names changed to the Australian Government, Department of Foreign Affairs and Trade (DFAT) and New Zealand Ministry of Foreign Affairs and Trade (MFAT) respectively. These RBM approaches were extended throughout these agencies with results aggregated from multiple programmes to report on progress within sectors and themes. Both agencies undertook a roll-out of these approaches with accompanying framework design and capability building programmes for managers and evaluators. A/NZ’s progress was outlined in a *Results in Development Cooperation Case Study* (OECD, 2017) which signalled positive changes from using results framework by staff and development partners.

Moreover, Picciotto (2009) outlined that the unit of analysis for development was increasingly moving to country and sector levels which included countries designing their own country, sector and agency strategic results frameworks, and to organisations and donor activities aligning to these frameworks at operational-level. This further shift led to an increased interest in the architecture and use of frameworks, different approaches for monitoring and evaluation, and new ways of thinking about management and governance systems in both developed countries and in developing countries. This

contextual shift was relevant to my case study research which focused on two national development, management and evaluations contexts, namely PNG and A/NZ.

## **2.6 Country systems**

The aid effectiveness principles (OECD2006) promoted harmonisation and alignment with national systems (particularly in those countries receiving donor aid) and the coordination of activities (national and development partner). Perhaps because of the relatively high profile of the Paris Declaration (Wood, Kabell, Sagasti, & Muwanga, 2008), the concept of country systems has emerged, and increased clarification of what this incorporates was needed. In commenting on the more bottom-up A/NZ response to the Declaration. Wood et al. (2008) observed that: “there is scope to enhance NZAID’s implementation of the Paris Declaration further, particularly alignment with country systems and processes, if the expertise of specialist contract and financial management staff could be harnessed more effectively” (p. 108). The rationale for an increased emphasis on country systems and accompanying strategic results frameworks was that in the previous decades, donors had by-passed country systems which can impact on the sustainability of national systems (OECD, 2009e).

In the *Accra Agenda for Action* (OECD, 2009b) “commitments to development alignment were re-iterated and deepened, and endorsed by over 100 countries and organisations... to enhance the mechanisms for mutual accountability” (p. 1). To enhance development effectiveness, this included amending existing approaches and practices such as increasing the alignment of activities by development partners with the partner country’s priorities and systems to eliminate duplication. Using country-led M&E systems incorporating results and outcomes frameworks aimed at enhancing evidence-based policy and decision-making. This increased focus on country systems led to new paradigms occurring in evaluation that focus on the principles of ownership and accountability. Ba Tall (2009) noted that “increasingly emphasis was being placed on strengthening country-led national-level M&E systems” (p. 1).

Segone (2009) surmised that an increase in country-led M&E systems and capacity would be required to provide better evidence of results as the knowledge base for country-led and owned M&E systems. Kindornay and Morton (2009) also advocated that “further studies with more consultation with southern’ stakeholders [that is, countries regarded as developing] may benefit this emerging knowledge area” (p. 1).

At the High-Level Forum on Aid Effectiveness in Busan, Korea in November 2011 discussions highlighted the need for a shift of focus “from effective aid to co-operation for effective development”(Busan High-Level Forum on Aid Effectiveness: Proceedings, p. 24). Furthermore, Betts and Wedgewood (2011) emphasised how “the Manila

*Statement on Effective States 2011* recognised the importance of results management and monitoring in reforming government and building effective and accountable institutions” (p. 8).

It is important that Development Partners support country capacity to mainstream results-based approaches into public sector management... there is a need for robust and demand-driven country-owned results-based public sector management reforms that build on existing country systems. (p. 8)

This increased focus on national and sector development is still emerging into mainstream development thinking. This re-emergent focus on national development was in response to “the challenges of how to manage the tensions inherent in processes of economic globalization, national sovereignty and democratic governance” (Rodrik, cited in Chimbhowu et al., 2019, p. 81). This shift of focus to national development appears to be contributing to an increased need for and role of evaluation at country and sector level, which was the focus of my research.

### **2.6.1. Governance and accountability**

Grindle (2007) outlined that increasingly governance was becoming a growing focus this century with increased capacities required of countries to govern effectively. According to (Kusek & Rist, 2004) it was widely accepted that “good governance is key to achieving sustainable economic development” (p. 1). However, according to Picciotto (2008), the accountability debate appears to focus predominantly on the programme level rather than on a sector or country levels which he considered was more important for analysis. Watts (2006) endorsed good governance and cited the World Bank’s dimensions of governance including: “voice and accountability, political stability, government effectiveness, regulatory quality, rule of law and control of corruption” (p. 607). Kaufmann and Kraay (2008) went further and included citizens’ views defining governance as:

...the traditions and institutions by which authority in a country is exercised. This includes the process by which governments are selected, monitored and replaced; the capacity of the government to effectively formulate and implement sound policies; and the respect of citizens and the state for the institutions that govern economic and social interactions among them. (p. 6)

Wagner (1989) outlined that there are several dimensions to accountability, including reporting and analysis, which need to provide an explanation and justification. Alkin (2004) considered that accountability also included answerability to stakeholders which was “not currently reflected in evaluation” (p. 14). Mayne and Rieper (cited in Gray et al.,

2003) supported this view for evaluation, saying that “public service quality is developed and implemented through tools in which evaluation plays, or should play, a significant role” (p.107). In his foreword to *Collaboration in Public Services* (Gray et al., 2003) Desautels, the Auditor General of Canada (1991–2001), reflected on the future role of evaluation and stated that:

Evaluation practitioners must wade carefully through the new governance models emerging. They must rise to the challenges in order to fulfil their responsibilities in support of accountability and transparency. (p. ix)

Schwandt (2005) emphasised that “practices are regarded as in need of repair by evaluation (and research) that can deliver the necessary science-based solution to the problems of practice” (p. 95). Stockman and Mayer (2016) in *The Future of Evaluation* emphasised the contribution that “evaluation can make to society”(p.10) particularly with social impact. This literature review highlighted multiple areas where evaluation could contribute to in an expanded role, and my case study research examined these areas and for other potential needs, particularly for strategic-level evaluation.

### **2.6.2. Nature of evaluation and implications**

The nature of evaluation has been the subject of much debate. Scriven (1991) regarded evaluation as “one of the most powerful and versatile of the trans-disciplines – tool disciplines such as logic, design and statistics – that apply across broad ranges of human investigation and creative effort while maintaining the autonomy of a discipline” (p. 1). Picciotto (2002) described how Campbell was one of the forerunners of evaluation where his “conception of the experimenting society raised expectations about the utility of evaluation for sound policy making” (p. 32).

Christie and Alkin (2003) questioned the use of the word “theory” within evaluation. They contended that it would be more appropriate to use the terms “approaches or models”, as none of the approaches described by evaluation theorists offers an empirical theory (p. 5). Christie and Alkin and considered that evaluation arose from two areas: “(i) the drive for accountability designed to improve and better programs and society; and (ii) social inquiry emanating from a desire for a systematic and justifiable set of methods for determining accountability” (p. 12).

One of the management challenges that evaluation has experienced was getting agreement on approaches and theory, given its interdisciplinary nature. Christie and Alkin and referred to those who have developed evaluation approaches and models as “evaluation theorists” and categorised the theorists’ views under three categories – “use, methods and valuing” (p. 3). Scriven (1991), a prominent evaluation theorist, emphasised

that evaluation involved making judgements on the “merit and worth” of programmes which he considers differentiates evaluation from research.

Evaluation theorists’ thinking relevant to results-based management was drawn predominantly from the theory-based evaluation authors as classified by Alkin (2004) under a “use” purpose. Broader discussions of theory-based evaluation were found in Weiss and Barton, 1980; Chen, 1990; Wholey, 1999; and Patton, 2002. Increasingly, since 2010, there was some agreement reached among the evaluation profession that outlining the theory of change from a programme and the use of frameworks to identify intended results were also referred to as “programme theory” (Funnell & Rogers, 2011; Rogers, 2009b). This approach to evaluation was in contrast to “goal-free” evaluation which was promoted by Scriven (1991) in his earlier work.

Since 2002, there has been an increasing shift in emphasis of evaluation as a function separate from management to a focus on “building internal evaluative management cultures and capability within organisations” (Ryan, 2003, p. 12) and Mayne (2007b). The United Nations Evaluation Group (UNEG) report *The Role of Evaluation in Results-Based Management* (2007) contained the findings and analysis of a survey of UNEG members to ascertain the role evaluation played in results-based management in various organisations. The key findings from this study were that the results-based management approach was seen primarily as a reporting tool rather than a management function, thus limiting the use of the information in decision-making.

United Nations Development Programme (UNDP, 2002) evaluation office views on the role of monitoring in the business of government were representative of the current approaches recommended by other evaluators such as (Rist, 2006). The response of UNDP to an intensified focus on outcomes was to advocate for “better measurement of performance and more systematic monitoring and reporting” (2002, p. iii) to foster an organisational culture of learning, transparency and accountability. Mackay (2006) identified three dimensions of success for development results: First, he suggested the utilisation of M&E information was necessary; and second, that good-quality M& E information was critical and, finally, that collecting such information should be sustainable. He further noted that benefits are only obtained when the results are used in budget, investment, and policy/strategy decision-making processes. He highlighted Chile as an example of what success looks like:

It is the intensive utilization of the monitoring information and evaluation findings which the M&E system produces ... [that] are used by the Finance Ministry for its resource allocation decisions with the budget process, and to impose

management and efficiency improvements on sector ministries in the programs for which they are responsible. (p. 3)

Mackay (2006) considered that M&E was essential to governance, not just to deliver outputs. Also, practitioners such as Rist (2006) and Stern (2008a) highlighted the need for evaluation to shift the focus to a more strategic level. Rist (2009) also considered that evaluation, as an operational practice, did not significantly include strategy and policy. With the drive for evidence-based decisions, it was becoming increasingly clear that there needed to be a paradigm shift in the role of evaluation.

More recent influences on how evaluation was framed and considered included indigenous evaluation approaches such as Kaupapa Māori evaluation (Carlson, Moewaka Barnes, & McCreanor, 2017; Cram, Pipi, & Paipa, 2018; *Enhancing Mātauranga Māori and Global-Indigenous-Knowledge*, New Zealand Qualifications Authority (NZQA), 2014) systems and complexity thinking and emphasis on the use of feedback loops for learning (Hummelbrunner & Jones, 2013; Williams & Hummelbrunner, 2011) and more integrated and adaptive participatory approaches (Shutt, 2016).

From the literature, more recent themes highlighted for evaluation to consider included increased stakeholder partnerships and participation which were valued by indigenous peoples, increased use of data-driven evidence for feedback and use in decision-making, expanded management capacity, and incorporation of a systems approach.

### **2.6.3. Performance management and role of evaluation**

Rist (2009) outlined that the growth of performance management in public sector governments expanded to meet the information needs of stakeholders. While performance management was a commonplace part of governance in development contexts (Mudacumura & Haque, 2004), the relationship between performance management and evaluation has remained unclear (McDavid & Hawthorne, 2006).

Multiple challenges in implementing results-based management approaches were outlined by the Auditor General's Office of Canada (2002), which highlighted how results-based management had proven to be quite challenging to implement. These challenges included the volume and relevancy of performance indicators, accessing results-based information, aggregating results from multiple programmes and the evaluative capacity and capability of managers. Moynihan (2005) noted that:

Although all levels of government have begun to pursue results-based reforms, which assume that managers will use performance information to make better decisions, reforms have neglected the insights of a large and relevant literature on organisational learning. (p. 203)

Ba Tall (2009) identified several issues in developing countries that were impacting on evaluation demand, including organisational thinking; learning capacity; and institutional weaknesses. She advocated identifying, creating and using knowledge while building capacity because of the significant growth in evaluative capability and capacity required. Hummelbrunner and Jones (2013) considered effective management required being more cognisant of contextual complexity, and they promoted using decentralised, collaborative and adaptive management tools.

Schwandt et al., (2016) emphasised there was a growing need for evaluation to provide more evidenced-based information for managers and at country-level to be able to assess progress towards the SDGs. More recently, the OECD (2019) published *Learning from Results-based management (RBM) Evaluations and Reviews* which highlighted that

Providers have made progress in integrating RBM in their internal systems (planning, implementation and reporting). As a result, they are better equipped to monitor and report on output data and short-term outcome data, and they are increasingly using results data for communication and domestic accountability purposes. (p. 5)

An increased emphasis on understanding the context and views of stakeholders was wanted, and adaptive management approaches advocated for at a corporate level included:

The essential elements of RBM systems such as a clear purpose for RBM, theories) of change, results frameworks, M&E processes, documented feedback loops and related decision-making and reporting processes, knowledge management and learning systems, as well as staff incentives. (p. 6)

Based on this review of literature, an expanding role for evaluation was noted mainly at the strategic level which appeared to require more adaptive management capacity and capability by managers and increased inclusion of stakeholders, and more transparent communication.

#### **2.6.4. Building evaluative approaches, capacity, capability, and culture**

Establishing an internal culture that is supportive of managing for results within institutions was emphasised in the previous discussion. Development of such an internal



culture relies in part on the use of feedback loops for learning and adaptive management as indicated by Argyris (1990a, 1997) and Schon (1963, 1983). They emphasised the importance of learning and using results-based data to inform decision-making within organisations. Segone (2009) outlined that if “evaluation systems are owned and led by concerned countries, it would facilitate the availability of more rigorous evidence such as country-specific data which is needed to monitor policy reforms and national development goals” (p. 23). He emphasised that from the Paris Declaration principles “the prominence of ownership reflects the understanding that national ownership and leadership is the most important overarching factor for ensuring good development outcomes “(p.26).

Moreover, Minkov & Hofstede (2012) recognised that national cultures have differences within population subgroups (such as cultural norms and values, gender, age) have an impact on different peoples’ needs, perspectives, values and behaviour within development and management practice. Hofstede (1991) “ranked countries on five value dimensions (1) power distance (2) individualism/collectivism, (3) masculinity/femininity, (4) uncertainty avoidance, and (5) short-term/long-term orientation” (p. 13). This rating aimed at deepening an understanding of how a country and its peoples’ multiple cultures operate. These dimensions were relevant to my research, for example as PNG tribes and Māori and Pacific peoples live predominantly in community and family settings and undertake more collective approaches to family wellbeing and prosperity which differs from European individualism. These differences impact on how different population groups within countries approach development, management and evaluate progress.

Schneider and Barsoux (1997) developed Hofstede’s argument further when they outlined how cultures also vary in terms of the way they establish truth and reality; for example, the relative importance placed on facts, figures and logic; as opposed to feelings, intuition and spirituality. Other cultural differences that impact within a country relevant to this research included tribal and patriarchal or matrilineal structures (such as in PNG). In A/NZ, the nature and role of evaluation still required consideration as there was diversity within and between the population subgroups including Māori iwi (tribes), European, Pacific, Asian and other groups. For example, Cram (2016) in discussing decolonising evaluation through using Kaupapa Māori approaches noted that there were multiple dimensions to attend to including the grounding and connections within a Māori worldview and social, economic and political accountabilities to their communities.

Other challenges for evaluation as noted earlier, included limited internationally agreed terminology in the results-based arena and few nationally mandated approaches

to designing, undertaking, and using evaluative information cognisant of a countries' population groups, cultures and values. However, countries increasingly were identifying and confirming their own approaches for evaluation which depend on the commitment to evaluation by individual governments, the stability of development objectives and access to sufficient resources.

Since 2010, some countries have moved towards a more systematic approach to evaluation, including Canada, USA, UK; South Africa and Uganda (Porter & Goldman, 2013). Gray and Bray (2019) in their paper on evaluation in the Australian public service reiterated similar issues to those highlighted in this section on the role and nature of evaluation at both the strategic and operational levels. They acknowledged that, internationally, other countries (such as US, UK, Canada and Netherlands) were shifting to more centralised and evidenced-based approaches and suggested measures that Australia could take to enhance knowledge and practice. From 2015, USA and UK governments have mandated roles for evaluation, and in 2016 PNG passed legislation for evaluation to be undertaken alongside their national development plan. In 2019, an approach and role for evaluation in A/NZ remained unclear.

## **2.7 Impact model**

I then designed an impact model (Figure 2) based on Brinkerhoff's (2002) *Success Case Study Method*. Key dimensions drawn from the literature were included in the impact model that contributed to "improved lives for people and the environment" as the ultimate outcome. The model suggested numerous opportunities for increased or improved processes and activities commensurate with the identification of good practice (Brinkerhoff, 2002).

The model dimensions were identified from engagement with the OECD aid effectiveness principles (2005a), development effectiveness and strategic architecture literature in development contexts, the use of results frameworks, initial stakeholder discussions about this research and associated good practice dimensions (drawn from literature and guidance) and intended impacts. These dimensions, included in the initial impact model, provided me with a baseline list (refer Appendix D, IV) against which I could systematically explore the relevancy and use of these dimensions from managers' and evaluators' perspectives. They assisted me to identify emergent findings and build theory.

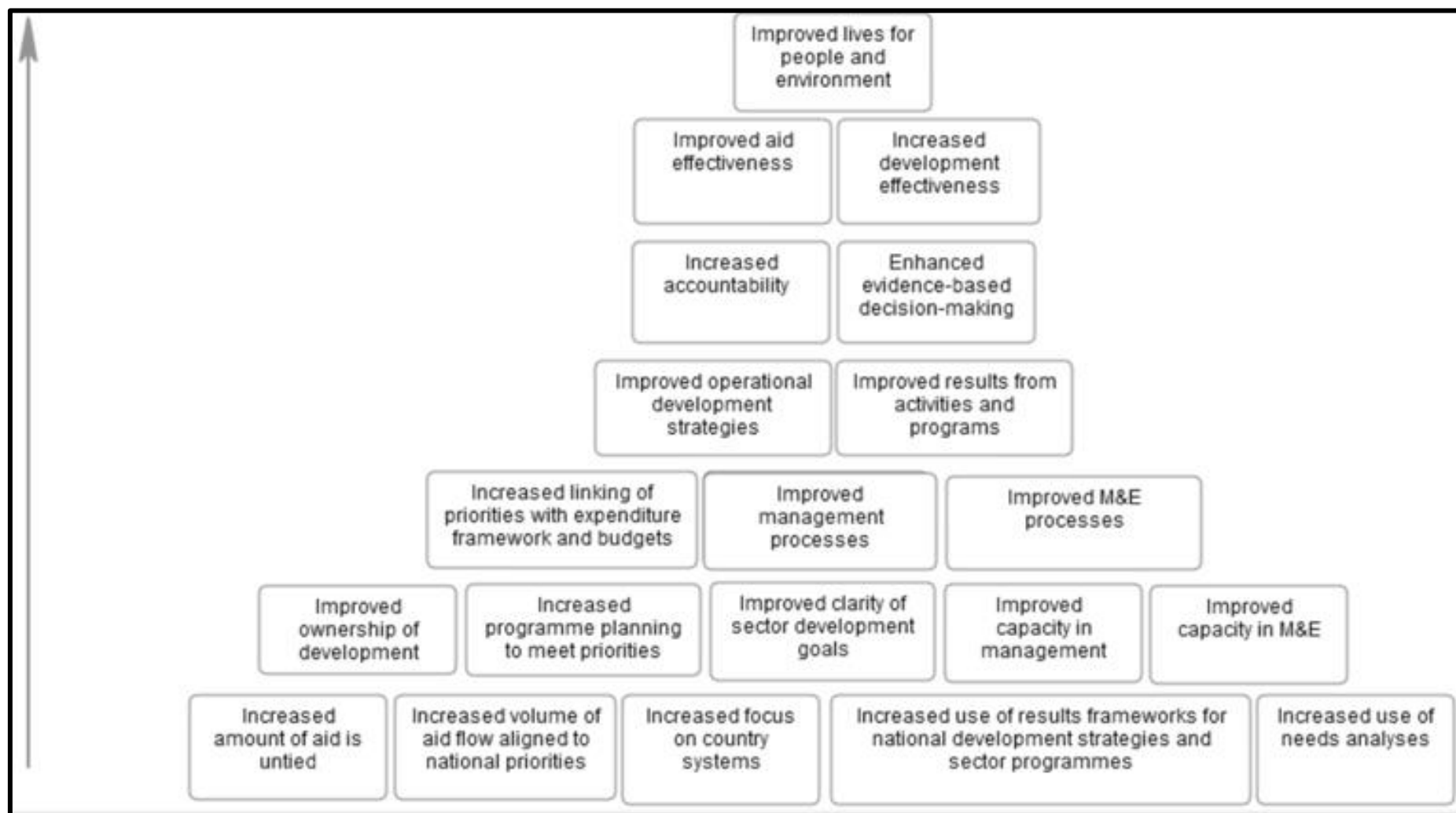


Figure 2. Initial research impact model.

The hunch I had, as an evaluation practitioner entering this research, was that there were likely to be more concepts to be cognisant of than were recognised in the literature and in practice because the dominant development paradigms eschewed values and or cross-cultural or non-western perspectives. A focus of my research was to identify good practice themes that could potentially be good practice concepts applied in different country contexts such as PNG and A/NZ. Further examination and a thematic analysis of the literature was undertaken iteratively during the writing up of this research aimed at making connections for more inclusive approaches to decision-making and learning, the architecture and use of frameworks, and the nature and role of evaluation within management and governance systems. The literature related to emergent good practice concepts identified from my research were discussed in Chapters Eight and Nine.

## 2.8 Themes

As I progressed with the literature review, I grouped common concepts that were emphasised under emergent themes. Each of these themes represented a gap in the ideas space around the development and use of strategic results frameworks as a tool at the crux of country, sector or even programme-led development practice, and within management and governance systems.

First, **the role and authority of strategic results frameworks** in the development and management context can be understood through a more critical lens. There appeared to be a genealogy of ideas that led to the idea of formalised strategic and business planning that has evolved from the new public management discourse. The influence of this discourse may need to be more clearly understood in terms of what is expected of national governments and donors responding to the OECD (2005a) commitments for development and aid effectiveness.

The use of frameworks was put forward as a tool for increasing harmonisation, alignment, managing for results and accountability between partner and donor countries (OECD, 2006a) without any explicit representation that those ideas stem from corporate goals for achieving growth within a Western economic development paradigm (Dwivedi et al., 2007). Extended use of strategic results frameworks may assist practice within national development planning to enhance country effectiveness and governance. However, more consideration appeared to be needed to encompass cultural values, principles and approaches to enhance inclusion, equity and sustainability. I argued that it was both the use of these frameworks with stakeholders and the way the development narrative was evolved and articulated within them that needed closer scrutiny.

There appeared to be limited development or articulation of underpinning values and high-level principles in the kinds of strategic results frameworks (country, sector, agency,

donor, results and performance) associated with management and governance systems in different development settings. There was also limited information on how to establish, measure the contribution of results from specific strategies, programmes and projects and aggregate results-based data towards a country, sector, and agencies' development goals cognisant of different population subgroups.

Second, the **inclusion of stakeholder perspectives** and involvement in strategic planning framework development and use has been given limited attention in existing discussions. I considered that frameworks, as artefacts produced in concrete form, maybe a useful tool to encourage dialogue around: priority-setting, evaluative planning; strategy development; budget alignment processes, implementation of activities; as well as the ongoing M&E of both outputs and development outcomes. These participative stakeholder and technical evaluative processes appeared interrelated and may contribute to a changing paradigm for evaluation. Chouinard (2013) outlined an emergent role for evaluation which could involve increasing links between participatory evaluation and accountability processes by extending stakeholder participation. I considered this research could contribute to evolving evaluation epistemology as it challenged the traditional view of evaluation focused on the merit and worth of programmes and projects being judged by evaluators as outlined by Scriven (1991, 1997) in his earlier views of evaluation. This research could contribute to an expanding narrative which increased value from "the logic of evaluation as outlined Scriven (2007). However, my research goes further to highlight the value of stakeholder perspectives at the strategic, sector, and country levels. -

Third, the role of **strategy and development planning** and the limitations of problem-based approaches within a construct of 'development and aid' can impact on development effectiveness. Focusing on "strategic intent" (Quesnel, 2007) was encouraged as an approach to incorporate more diverse population group views within a development context and enhance ownership by country stakeholders. Furthermore, Pacific cultures advocate for increased use of strengths-based approaches (as opposed to problem or deficit-based) to improve relevancy and appropriateness for country, sector and regional strategic development in different contexts. These changes included between (1) developing countries and their development partners, and (2) indigenous and specific population groups within country development contexts where cultural values and norms differ from Western values and norms. An example of these differences was in A/NZ where Māori and Pacific peoples culturally value more collective, spiritual and family-based approaches (Cram, Pipi, & Paipa, 2019; Ministry for Pacific Peoples, 2017) which differed from a European focus on individualism. My research aimed at contributing to this ideas space by examining emergent concepts that could be

relevant and applicable within different contexts for country development and their diverse population subgroups.

Fourth, the **nature and role of evaluation and governance** in development management was another knowledge space that required further examination. An increased role of governance was advocated for by Dormer and Ward (2018) to enhance “accountability for outcomes (or results)” (p. 2). There also appeared to be capability, and institutional exogenous and endogenous challenges (OECD, 2019) noted with results-based management and the role of evaluation with underpinning assumptions evident such as “power and relationships and capacities” (p. 28) which can impact on approaches for development management. This research has focused on drawing out potential good practice concepts that may enhance awareness of and development management knowledge and practice within different contexts.

Fifth, the relevancy and **underlying mechanisms** of potential good practice concepts were knowledge ideas that required further examination within the contextual layers (such as *wider infrastructure*; *institutions*; and *ideas, interventions and frameworks*), which were initially adapted from Pawson’s layered model (2006, p. 18). The role and authority of the architecture and use of frameworks within development management systems were central to this research. It appeared that developing an effective architecture and use of frameworks relies on individual capacities, interpersonal relations and institution and wider infrastructure settings which may need to be enabled through approaches established within a national context. How national governments, development institutions, and other stakeholders engaged in these approaches was examined in my primary research.

Sixth, while undertaking this research from 2012 onwards, the paradigms for development shifted further with the introduction of the Sustainable Development Agenda 2030 (UNDP, 2015b). This research was then emergently framed by a shifting context calling for more **sustainable development and adaptive management approaches**. An increased role for governance and a need for more inclusive relationships has increased over the past five years. A growing emphasis on economic, social, and environmental development dimensions, particularly with climate change and social and regional inequalities have emerged which require more explicit consideration in both developing and developed countries. This research explored how sustainability, inclusion and adaptation concepts could be applied within a country development context.

## 2.9 Knowledge space for this research

My research set out to examine the changes in development and evaluation practice relevant to the architecture and use of strategic frameworks, management and governance systems. The social reality layers of *ideas, institution, and wider infrastructure* were explored, alongside the implications for practitioners, management and governance in terms of evaluation. The literature on development, evaluation, management, governance and the history of, and knowledge underpinning the architecture and use of frameworks revealed gaps in the knowledge underpinning the development and use of strategic frameworks, management and governance systems.

The literature reviewed suggested that any crossover of knowledge between evaluation practices in different development and evaluation contexts (that is, developed and developing countries) was limited. For example, the knowledge concepts used in evaluation in developing countries and international development had not drawn at that time on some key evaluation practice elements such as program theory (Rogers et al., 2000) which was becoming more used in developed countries such as Australia. Conversely, strategic results frameworks were becoming an essential M&E tool used in international development yet were as not commonly used in developed country contexts. Moreover, how these evaluative tools were developed and used in either contexts required further examination. Changes were evident with some crossover and knowledge with concepts such as theories of change (Funnell & Rogers, 2011) and results frameworks introduced within different development contexts such as Australian Department of Foreign Affairs and Trade (DFAT), 2017; New Zealand Ministry of Foreign Affairs and Trade, 2016). These concepts superseded the logical framework approach. However, I noted that a knowledge space remained between strategic management and evaluation concepts which were the focus for my research.

My research included examining links between planning for and measuring outcomes and use of evaluative information in decision-making by managers within management and governance systems in different country contexts. My literature review also highlighted that the impacts and significance of changes on evaluation and management knowledge and practice needed to expand to include aspects of governance. This change was due to the wide-ranging impacts from using strategic results frameworks within development settings which goes beyond the management of specific activities or programmes and included stakeholder engagement and perspectives.

From my literature review I found there appeared to be is a growing interest in increased knowledge and use of values and principles with country and sector strategic results frameworks and for strategic evaluation approaches within management and governance systems. The overarching research question for this thesis was *What 'good*

*practice' evaluative principles and concepts may enhance country, sector and agency development, management and governance knowledge and practice in different contexts?* Thus, one of the directions for this research was to identify and examine potential values and principles and assess their relevancy and impact for managers, evaluators, management and governance systems, which could, in turn, enhance inclusion, development effectiveness, governance and sustainability. Chapter Three provides a discussion of the two case study research contexts (A/NZ and PNG).



## Chapter 3: Methodology

The methodology chapter structure follows with a discussion of the theoretical perspective underpinning the research. Second, the research design and methodological considerations are highlighted. Then the research methods and data collection tools are outlined.

### 3.1 Theoretical perspective and research frame

This research involved the use of theory or logic to depict reality and the critical examination of how that operates in the “real world” (Bamberger et al., 2006). Critical realism (Archer, Bhaskar, Collier, Lawson, & Norrie (eds.), 1998) was selected as the theoretical perspective underpinning this research to gain a more in-depth understanding of the architecture and use of results frameworks, and the interface between development, management, evaluation and governance knowledge concepts. Wuisman (2005) considered that “the critical realist approach to social scientific research starts from the ontological notion that social reality is stratified”:

A central idea of critical realism is that natural (physical and biological) and social sociological reality should be understood as an open stratified layered system of objects with causal powers making things happen. (p. 368)

Bhaskar (Archer et al., 1998) distinguished three hierarchically arranged layers, namely the empirical (experiences), the actual (events) and the real (mechanisms).

Layers of reality	Positivist/ empiricist approach	Interpretive constructivist approach	Critical realist approach
1	Empirically observed characteristics	Symbolically expressed meanings	The empirical (experiences)
2			The actual (events)
3			The real (mechanisms)

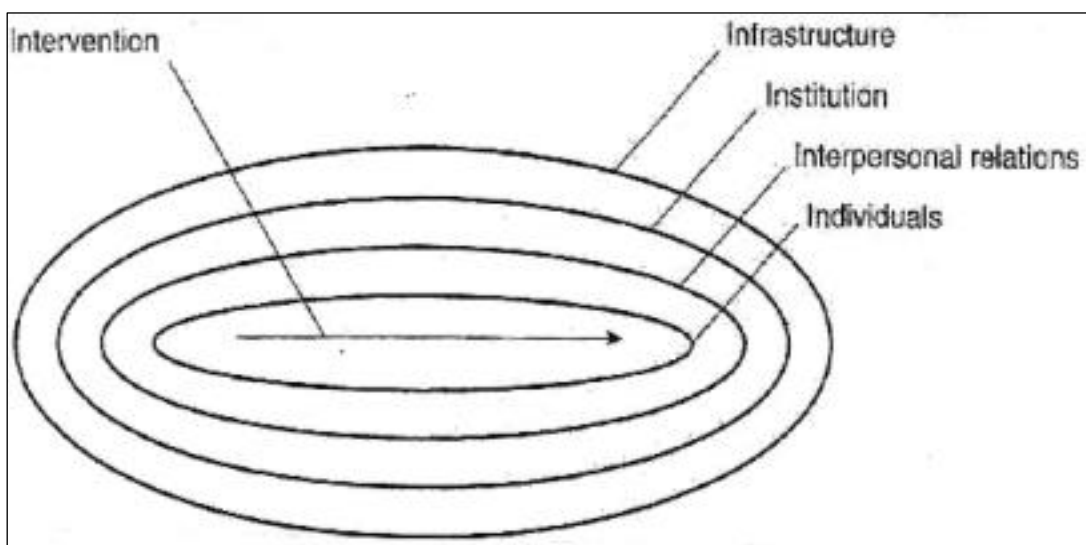
**Figure 3. Ontological conception of social reality (Bhaskar, 1979/1998, p. 51).**

This research initially set out to use Bhaskar’s idea of social reality stratification layers as an ontology (Wuisman, 2005) to assist in researching, analysing and identifying good

practice principles (conceptual and operational) to potentially extend knowledge and support enhanced practice. However, unanticipated findings emerged during the scoping phase from the key informant interviews (including managers and evaluators/performance management specialists) that pointed to additional relevant dimensions or concepts as being important over and above those initially incorporated in the research impact model (Figure 2).

After further analysis and reflection on themes emerging from my diagrams (displayed in Appendices A & B), as part of my theory-building research approach, I felt these unanticipated findings could be grouped under *ideas, interventions and frameworks; institution; and wider infrastructure* settings. Using “pattern codes and constructs” (Miles & Huberman, 1994, p. 69) in this somewhat complex context of country development, governance, management and evaluation allowed me to sense-make and group concepts and themes based on Pawson’s (2006) contextual layers.

Therefore, rather than adopting Bhaskar’s approach at this stage, I noted Pawson’s (2006) discussion of how causation in the social world could be construed through his classification of “five contextual layers – ideas or interventions, individual capacities, interpersonal relations, institution and infrastructure” (p. 31) (Figure 4).

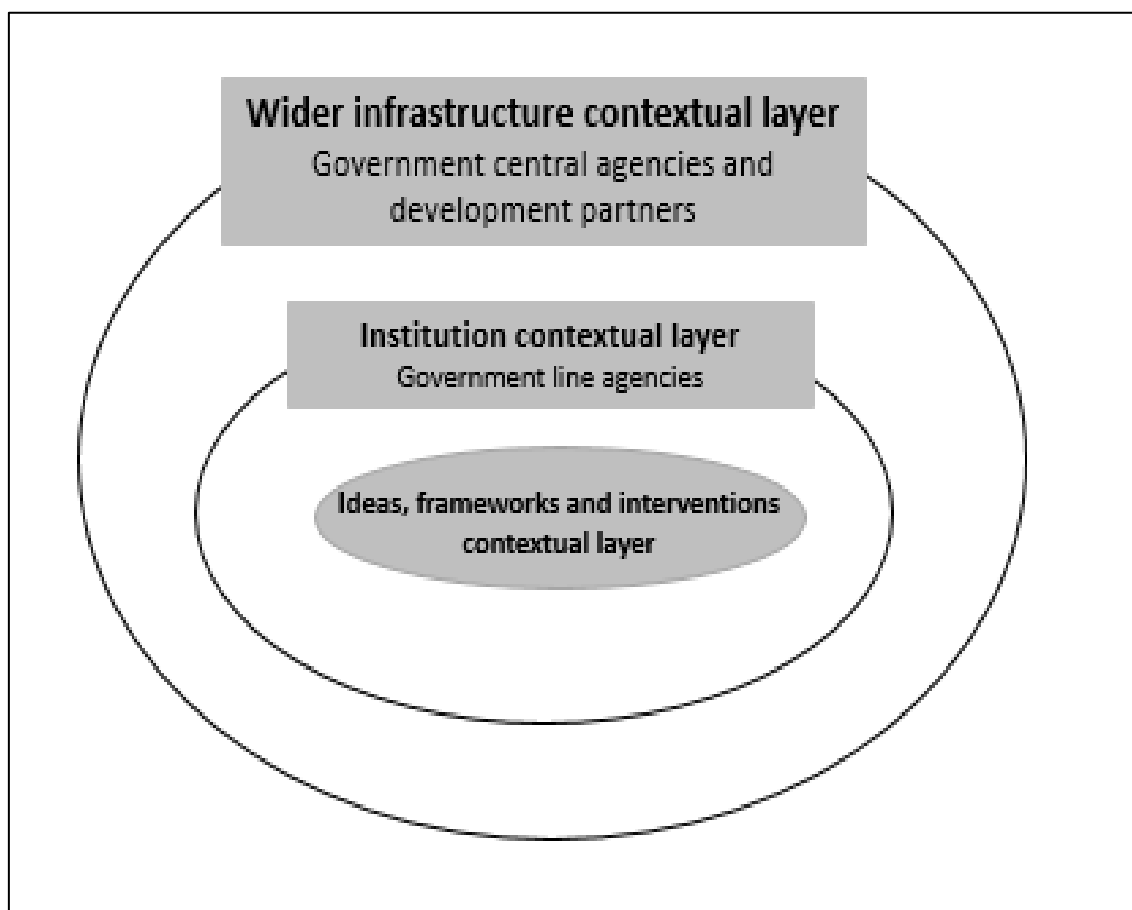


**Figure 4. Pawson’s hierarchy of contextual layers showing “the intervention is the product of its context” (2006, p. 31).**

I considered that “ideas and interventions” as depicted by Pawson (2006, p. xiii) were displayed as (results) frameworks for agency and sector business as usual activities and interventions. Pawson referred to strategies and programmes as “embedded theories in active and open systems” (p. xiii).

The final research frame incorporated an adaptation of Pawson’s (2006) classification (Figure 4) as an alternate, new typology for grouping concepts within the contextual

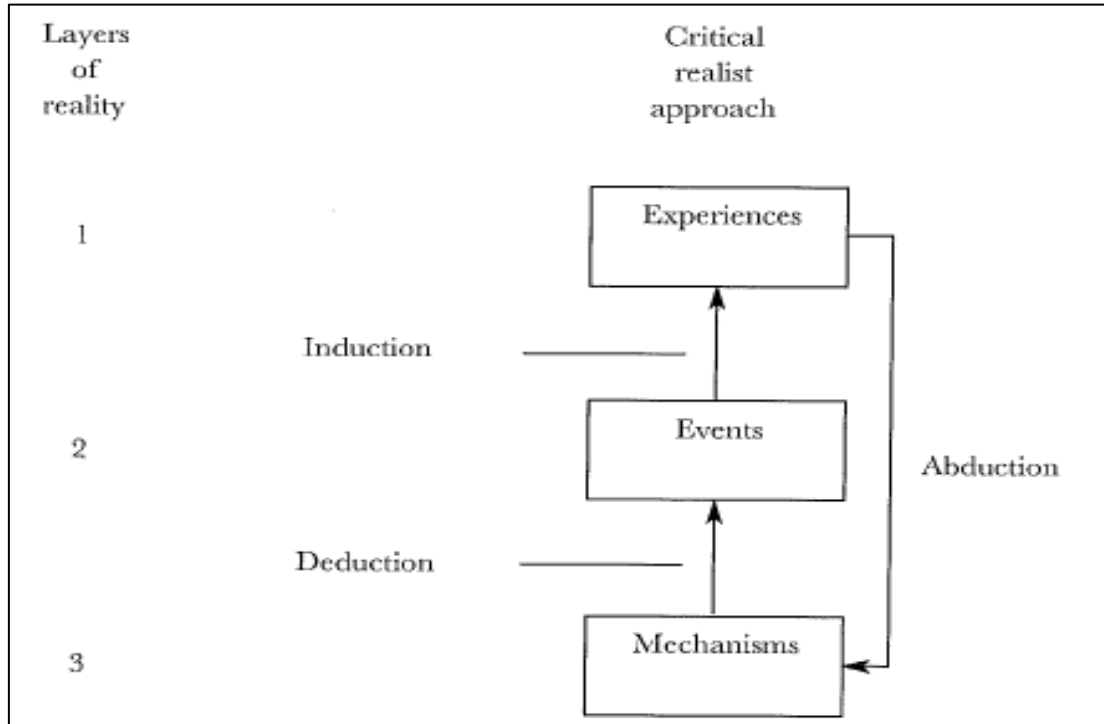
layers. I developed my own typology based on the scoping phase research findings, my practitioner 'lived' experiences and academic studies, as I considered that Pawson's contextual layers were not a 'perfect fit' to use in my case study research (Phase Two) and analysis. What I observed from my findings was that dimensions of human agency (such as the individual 'capacity' of managers and evaluators and their interpersonal relations) appeared to be operating differently within each contextual layer. This epistemological reflection proved critical for reframing my conceptual research frame (Figure 5) for the two country case studies. This led to my research design being adapted between the scoping phase and case study research. During the scoping phase of my theory-building research, I developed a new conceptual research typology, which was drawn from the emerging research findings from the scoping phase and my 'lived' experiences as an evaluation practitioner. I used this new conceptual research typology throughout my case study research and primary analysis phases, which assisted me to structure my findings, theory-building thinking, discussion and conclusions.



**Figure 5. Emergent conceptual research typology and primary analytical frame.**

This led to me reconsidering how I could analyse the emerging themes within each contextual layer as a secondary analysis to examine how these themes operated more deeply. Wuisman (2005), in his discussion of Bhaskar's diagram (Figure 6) on the logic

of scientific discovery considered that induction, deduction and abduction all play a role to “explicate the line of argument involved in the ontology of critical realism” (p. 384). He considered that abduction “helps specify the relationships of causal necessity between the different ontological levels” (p. 384).



**Figure 6. Layers of reality and the role of abduction (Wuisman, 2005, p. 384).**

Critical realism combines ontological realism (Wuisman, 2005; Maxwell & Mittapali, 2007) with epistemological constructivism using deduction and abduction to identify underlying mechanisms from the events and experiences. Realism looks at explanations for real events and assists with causal explanations. Morton (2006) emphasised examining the three strata: experiences, events and mechanisms. The empirical were the observed experiences; the events were actual events generated by the mechanisms, and the real mechanisms were the structures and processes that generate events. The criticism within realism by researchers where they consider criticism is warranted is based on their understanding of the particular social system (Archer et al., 1998).

Development, management and evaluation systems (Dalkin. Greenhalgh, Jones, Cunningham & Lhusser, (2015) incorporate the architecture and use of results frameworks in specific contexts and the interface between people. Critical realism also enabled me to explore different perspectives of stakeholders within the ontological contextual layers (Pawson, 2006). For me as the researcher, I considered the nature and the perspective of the criticism required examination within country development, management and evaluation contexts to understand what important concepts for managers and evaluators were that may enhance development effectiveness.

Critical realism theorists debated whether thoughts and actions can be subject to criticism from a particular view (Bhaskar, 2002; Pawson, 2013). Hammersley (2009) noted that “even where value judgements rely on research evidence, they also necessarily depend upon other factual assumptions, and upon value principles that are plural and often in conflict”(p. 7).

My research examined development effectiveness and the role of governance within national contexts. It included how different population subgroups’ development goals (Dwivedi et al., 2007) were considered. I wanted to gain an understanding of who was involved in development planning cognisant of location, ethnicity, and culture, and in making evaluative judgements (Scriven, 1997). Hammersley’s (2009) view of criticism within realism had relevance to this research when considering the different perspectives and values of the stakeholders (national governments – managers and evaluation specialists and development partners) which underpin the architecture and use of strategic results frameworks in different contexts (for this research it was country systems and approaches for development, governance, management and evaluation in A/NZ and PNG). I considered it was important to examine the different perspectives of and values underpinning the contexts and stakeholders’ knowledge and practice in this research.

In my research, I used a critical realism ontology in two ways: first by incorporating three contextual layers (*ideas, institution and wider* using the strata of experiences, events and mechanisms as a secondary analytical frame within each social reality layer. These different layers and strata assisted in identifying nested good practice concepts. By applying an ontological conception to my research, I realised that these embedded theories could be development narratives depicted as country, sector or agency strategic results frameworks and strategies with associated interventions and theories of change as outlined in the OECD (2006a) guidance. Results-based management and theory-based evaluation (Chen & Rossi, 1980) were premised on an idea, theory, or intervention which was then implemented, monitored and evaluated, within *institution* and *wider infrastructure* settings.

After my fieldwork research was completed in these two country contexts, I undertook a secondary analysis of the research findings. This included the critical realism strata of experiences, events, and mechanisms as outlined by Bhaskar (1988) and Wuisman (2005) within each of the contextual layers. My scoping phase research showed that interpersonal relations and individual capacities appeared to operate within each of the three contextual layers of (1) *ideas, interventions and frameworks*; (2) *institution* (or line agencies) ; and (3) *wider infrastructure* settings.

Overall, I found this combined use of contextual layers combined with a critical realism strata enabled me to examine what was happening more deeply. The identifiable key concepts from my research were discussed in Chapters Eight and Nine, after the findings for each of the three layers which were outlined in Chapters Five, Six and Seven.

### **3.2 Research questions**

My thesis was that current evaluative approaches to development, management, and governance, particularly in the context of the SDGs, were insufficient to ensure robust, inclusive, and effective results. The overarching research question that I focused on to explore this argument was: *What are 'good practice' evaluative principles and concepts that may enhance country, sector and agency development, management and governance knowledge and practice in different contexts?*

The refined sub-research questions were part of an emergent research design:

1. What evidence from A/NZ and PNG demonstrates what works for whom in relation to results frameworks and associated management and governance systems? (RQ1) (Chapters 5, 6, and 7)
2. How and in what ways can conceptual links be identified between frameworks, management and governance systems and good practice principles? (RQ2) (Chapter 8)
3. How and in what ways can country, sector, and agency strategic results frameworks and their associated management and governance systems underpin the development of good practice values and principles (RQ3) (Chapters 8 and 9)
4. How do the emerging good practice principles in country, sector, agency and programme systems work to enhance development effectiveness and governance for development (national governments and partners), management, and evaluation? (RQ4) (Chapters 8, 9, and 10)

Primary research encompassed a wider analysis and two comparative case studies (one deeper level which included more interviews and a rating table completed by research participants in A/NZ and one lighter level of research in PNG). A complex and open research design led to several unanticipated findings. After an initial scoping phase, I used an evaluative case study approach as a method, in the context of a critical realist perspective that was abductive and emergent. Low-risk ethical approval was given for by Massey University based on the information provided (see Appendix C).

### **3.3 Research methodology**

The research was undertaken in two phases where initial scoping preceded fieldwork on the two evaluative case studies. The impact model (Figure 2) drawn from the OECD (2005) guidance and good practice concepts identified during the literature review. During the key informant interview scoping phase of the research, an inductive approach (Miles & Huberman, 1994) was used to assist with exploration, discovery, and inductive logic for the scoping phase interviews (Appendix D, interview guide 1). This phase also included a scan of operational material (performance management guidelines, M&E guidelines, budget cycle, etc.) to understand the contexts for this research.

This initial literature search and reading pointed to the contestation of terminology (and therefore understanding) as well as a gap in the body of knowledge relating to concepts underpinning the architecture and use of strategic results frameworks, evaluative capability and capacity constraints, and limited use of results frameworks by managers and evaluation practitioners, institutional capacity and capability issues, and broader contextual influences. I noted that the earlier literature focus on frameworks appeared to be predominantly at the operational-level (results from programmes and projects) and to a more limited degree on the architecture of frameworks at strategic-level. Moreover, the literature on the interface between strategy, management and the role evaluation in different development settings appeared limited, which may be why I experienced both knowledge and practice gaps within multiple development settings.

Phase 1 of my research pointed to the importance of contextual layers (such as individual capacities, interpersonal relations, institution, and wider infrastructure settings) as identified by Pawson (2006) in which frameworks are nested. Based on these emergent findings from the scoping phase, a fieldwork phase (Phase Two) was developed. This second research phase incorporated open questions with both management and evaluation stakeholders in two development contexts (A/NZ and PNG).

The methodology for the thesis was emergent in the sense that, while interpretivist grounded theory (Oliver, 2011) and applied evaluation perspectives (Patton, 2002) provided one path into a research design, a critical realist approach (Wuisman, 2005) provided another. I explored both methodological options from the outset. It was unclear initially how and in what ways these approaches could work together or complement each other. A reflexive approach (Alvesson & Skoldberg, 2009) provided a way for me to capture a picture of the methodological complexity at the scoping stage.

### **3.3.1. Emergent research design frame – case study**

As I examined the initial findings from the scoping phase interviews which research participants considered underpinned the architecture and use of frameworks, it was apparent that managers and evaluators did not share a common understanding of concepts displayed in the impact model. I found that by using an updated interview guide (Appendix D, interview guide 2) with more open-ended interview questions utilising aspects of grounded theory led to broader knowledge and practice concept were identified by stakeholders. I found that these emerging concepts and findings emphasised by managers and evaluators were more comprehensive than the OECD guidance (2006).

From my analysis of the findings identified during the initial key informant interviews in the scoping phase that the knowledge gaps in the literature review were also apparent in practice. This realisation led to a change in research design. I considered that continuing to use only a theory-based frame for the research based on the impact model might limit data collection and the possibility of uncovering further unanticipated findings. By applying aspects of grounded theory (Oliver, 2011), it was possible to use more open questions in the evaluative case study interviews. This approach enabled me to draw out emerging themes in national planning, management and evaluation systems from different perspectives in different settings.

#### ***Use of aspects of grounded theory***

Adapting grounded theory within a critical realism paradigm was advocated for by Oliver (2011). She considered that critical realism and grounded theory “share a focus on abduction” (p. 1) and together, can provide a robust research approach with an “interconnectedness of practical and theory” (p. 1).

The combination of grounded theory within a critical realism paradigm in my research led to (i) use of open questions for interviewing and in the analysis (ii) a process of emergent coding, and (iii) an analysis of the emergent findings in the case studies. Dey (1999) outlined that in this:

process of emergence, the researchers have to rely on their own ‘theoretical sensitivity’ to generate relevant categories from the data. The researcher has to be able to think theoretically – to glean insights from the evidence, to conceptualise the data, and then to analyse relationships between the concepts.  
(p. 4)

I considered this emergent research approach allowed for a more naturalistic strategy for data collection in the latter part of scoping phase and for the case studies. It assisted me to avoid being captured by preconceived ideas. Data was collected using the open



questions in interviews with managers and evaluators and coded under emerging themes. For the initial interviews for the scoping phase, I used Interview Guide 1 (Appendix D) based on the impact model. These initial interviews were recoded using an open coding technique (Glaser & Strauss, 1967) to document emerging themes.

Based on these emergent findings from the recoded interviews in the scoping phase, a more extensive research design was used for the case study research. This incorporated open questions with both management and evaluation stakeholders in two development contexts (A/NZ and PNG).

### **3.4 Research phases and methods**

This research incorporated two phases with associated research methods:

1. **Scoping phase** – key informant interviews (18) to gain context, perspectives and identify knowledge and practice concepts. This phase resulted in an emergent research design using elements of grounded theory due to unanticipated findings beyond the dimensions included in the impact model. My initial impact model (Figure 2) was adapted to include the three contextual layers of the emergent research conceptual typology (Figure 7) for the case study research in Phase Two.
2. **Phase Two:** Case study research in A/NZ and PNG included:
  - i. Qualitative interviews (36) with managers and evaluation practitioners from central and line agencies and funders in both PNG (17) and A/NZ (19) contexts. The interview guide used for case study interviews incorporated more open-ended questions to draw out the extended concepts from the research participants. Deeper research was undertaken in A/NZ with additional interviews and a more in-depth review of the accountability documents. The lighter level was due to the availability of research participants and more limited availability of line agency accountability documents.
  - ii. Rating of impact model dimensions was completed by A/NZ case study research participants.

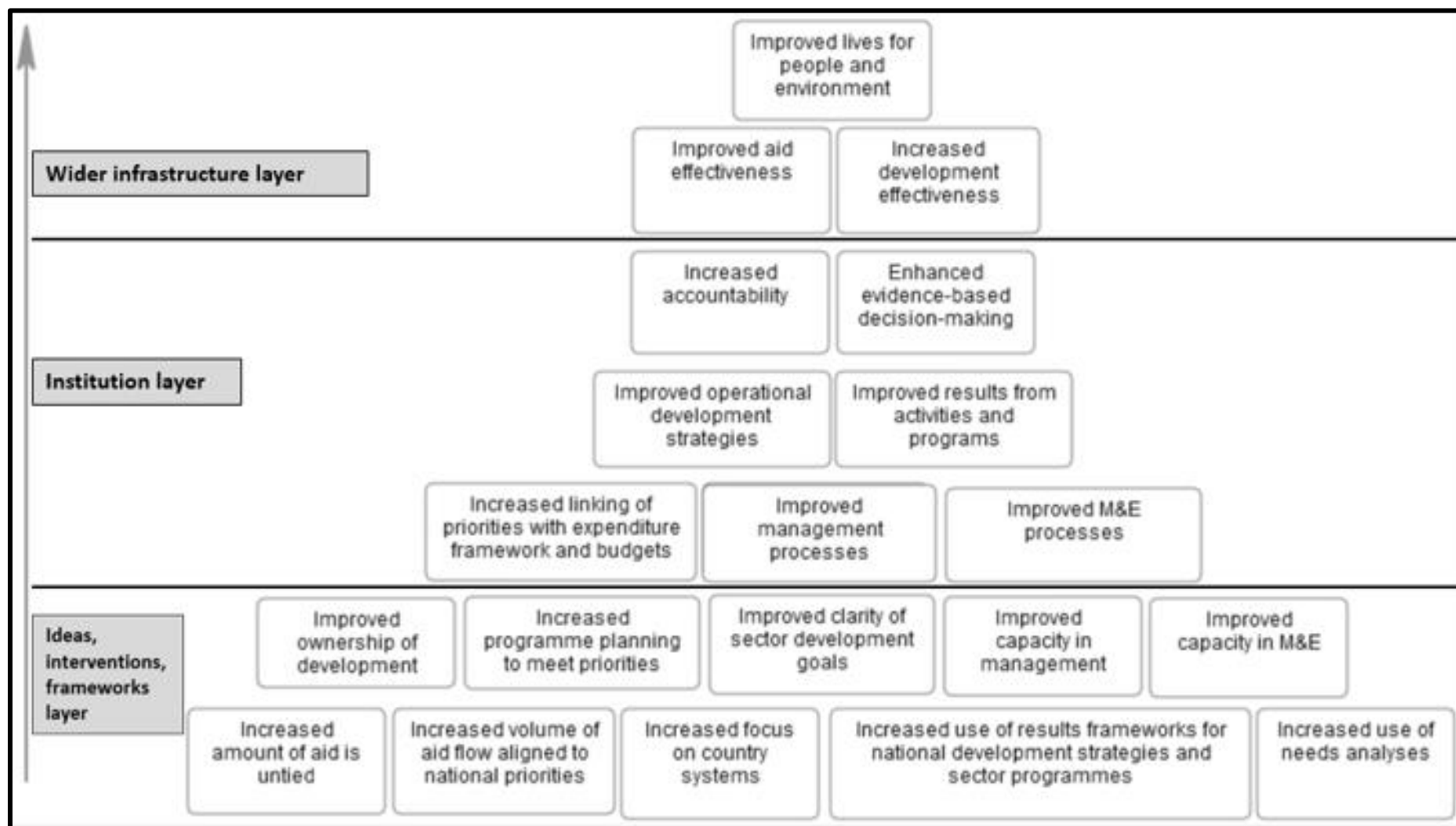


Figure 7. Updated impact model – initial research frame (superimposed as an emergent research frame).

### **3.4.1. Scoping phase: understanding the topic**

The aim of the primary research scoping phase was to understand the broader context and current thinking underpinning the architecture and use of strategic results frameworks, and what was changing within management and governance systems. The scoping phase provided the broader context, the logic of inquiry and confirmed the typology of the case study research in two different country systems.

First, a scan of guidance documents was undertaken on the current approaches and practice with results frameworks at both strategic and operational levels. The rationale for looking at both levels at this stage was to understand the context and the links between the strategic and operational level use of results frameworks. The detailed use of results frameworks at programme and project levels was not included within the scope of this research. The documents included guidance from country systems in A/NZ, Australia, Canada, the UK, the USA, Ghana, PNG, Samoa, and South Africa and donor guidance including A/NZ Ministry of Foreign Affairs (MFAT, formerly NZAID), Australia Government Aid Programme (DFAT formerly AusAID), Department for International Development (United Kingdom), OECD, European Union, United Nations Development Programme, World Bank, and Asian Development Bank.

#### ***Key informant interviews***

An interview guide (Appendix D interview guide 1) was developed for the scoping phase based on the impact model. The key informant interview participants were identified using a purposeful sampling approach (Patton, 2002). This sampling included snowball sampling (Miles & Huberman, 1994) for interviewing with senior managers and evaluation specialists identified in agencies and roles by positional and knowledge specialists.

Eighteen interviews were undertaken at this initial stage using informed consent processes. These interviews (face to face and telephone) included:

- i. A/NZ (Treasury, SSC, Office of the Auditor General, and a limited sample from public sector agencies).
- ii. PNG (Office of Auditor General, DNPM, Department of Provincial Local Government Affairs).
- iii. Donors (NZAID, AusAID, Asian Development Bank, UN agencies, The World Bank).

These interviews provided background and central and agency-based information and data, which assisted with structuring the case study research. The interviews were digitally recorded and transcribed to enable coding under themes (code frames), in NVivo

– specialised qualitative software was used to analyse and manage the data (Bazeley, 2007).

An accountability document review template was developed framed around the good practice concepts identified during the literature review (see Appendix D). A five-scale rating system was utilised for the document analysis of A/NZ government agency accountability documents – Statement of Intent (SOI) and annual report. These documents met A/NZ line agency mandated annual accountability requirements. The alignment of agency strategic plans was reviewed where available. These documents were referred to as the key accountability documents in the central agency guidance to A/NZ agencies (SSC, 2006).

### **3.4.2. Phase 2 – Case study research**

An evaluative case study method was used “for learning about a complex instance, based on a comprehensive understanding of that instance obtained through extensive description and analysis of that instance taken as a whole and in its context” (Morra & Friedlander, 1999, p. 3). Case studies were selected as an appropriate method for this research to gain this conceptual information. The selection of case studies was congruent with Yin’s (2009) view of selecting the case study research as a preferred method when “(a) how or why questions are being posed, (b) the investigator has little control over events, and (c) the focus is on a contemporary phenomenon in a real-life context” (p. 2).

All of these points highlighted above applied for this research as (i) frameworks were imposed by key stakeholders within country and sector settings, (ii) I wanted to gain additional knowledge in the associated management and governance systems and therefore the ‘how’ and ‘why’ questions were important; and (iii) I wanted to know what was happening, how it worked in different contexts, and what was planned or thought about going forward.

#### ***Case study criteria***

Interest was displayed for the case study research by governments in A/NZ, Australia, Solomon Islands, PNG, Samoa, Cambodia and Laos from visits to different countries in the Pacific and South-East Asia. Discussions with key stakeholders ranged from Chambers of Commerce, government ministers, donors, national organisations and government departments. The location of the two selected case-studies was based on context, development and evaluation progress, and access to information, within resource constraints. The case studies were selected based on the following criteria:

1. A/NZ: developed country, researcher's home country, government guidance incorporated a results-based public sector country approach using outcome frameworks at sector/agency level. Central agencies provided limited operational guidance to agencies due to devolved management to chief executives. A/NZ was a member of the OECD and had agreed to OECD-DAC Aid Effectiveness Principles.
2. PNG: had agreed to the OECD-DAC (2005) Aid Effectiveness principles, was a member of the OECD and had an emerging national evaluation community. GoPNG utilised centralised planning approaches with an application of a country and sector frameworks and were evolving their linking between development planning and evaluation.

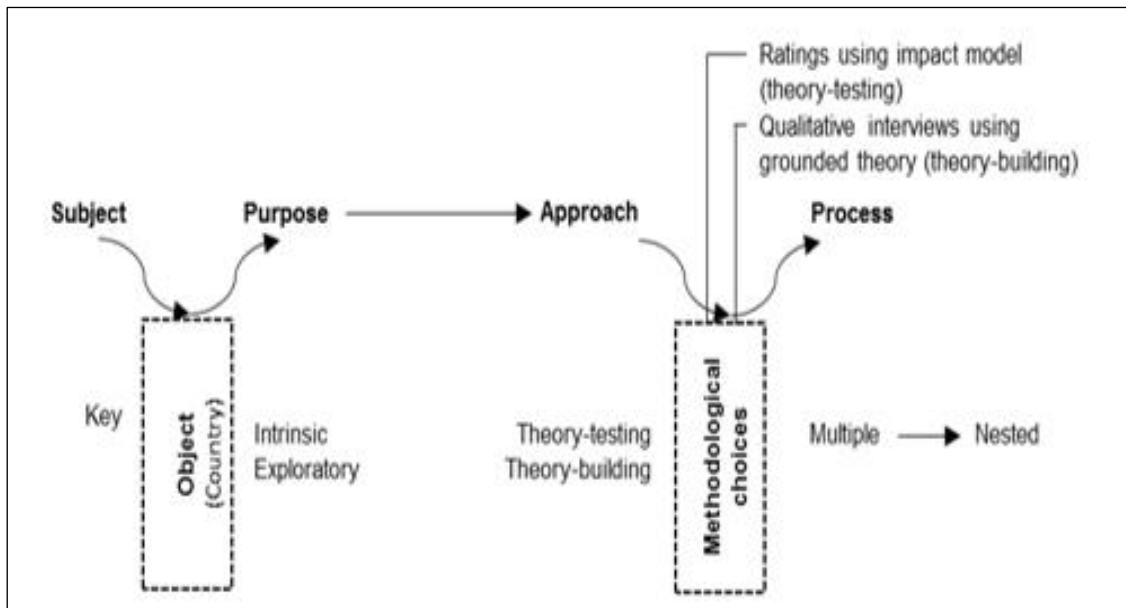
The first case study undertaken was at a deeper level for the A/NZ country system due to proximity, availability and focus of the study. The multiple research methods included scoping phase key informant interviews from A/NZ central and PNG agencies (18 interviews), a document scan of central agency planning, performance and accountability guidance documents, a rating review of accountability documents (SOIs and annual reports) using a rating matrix-based from impact model and literature review themes, individual interviews with both A/NZ managers (12) and evaluation /performance measurement specialists (7) including up to six line and two cultural agencies, and interviewees provided a rating from agencies on impacts from taken directly from the impact model (using IV Case study – impact model result rating table, Appendix D).

The PNG case study interviews (17) included central agencies (6) and line agencies managers (4) and evaluators (4) and donors/development partners (3). The donor/development partners included both Australian and A/NZ Foreign Affairs and Trade and development aid staff (managers and evaluators).

In addition, three A/NZ interview participants had worked previously for the World Bank, Asian Development Bank and United Nations Development Programme.

### **3.4.3. Case study typology**

A case study structure was required to undertake the research systematically across the two case studies. Thomas (2011) outlined that the case study contained two elements: "(i) subject of the case study – 'practical, historical unity,' and (ii) the object of the study – analytical frame" (p. 511). Figure 8 demonstrates how Thomas's typology of the case study research applied to this research.



**Figure 8. The typology of the country case study research and associated research methods.**

In this research, the case study subject was the country. The subject of the case studies selected were the changes underpinning the architecture and use of strategic results frameworks and associated management and governance systems in the context of countries and donors/development partners in the Pacific. The country-level approach to the architecture and use of frameworks, and the impacts and significance of the changes in terms of development (national governments, and donors for PNG), management, evaluation and governance in two different country settings (A/NZ and PNG) were examined.

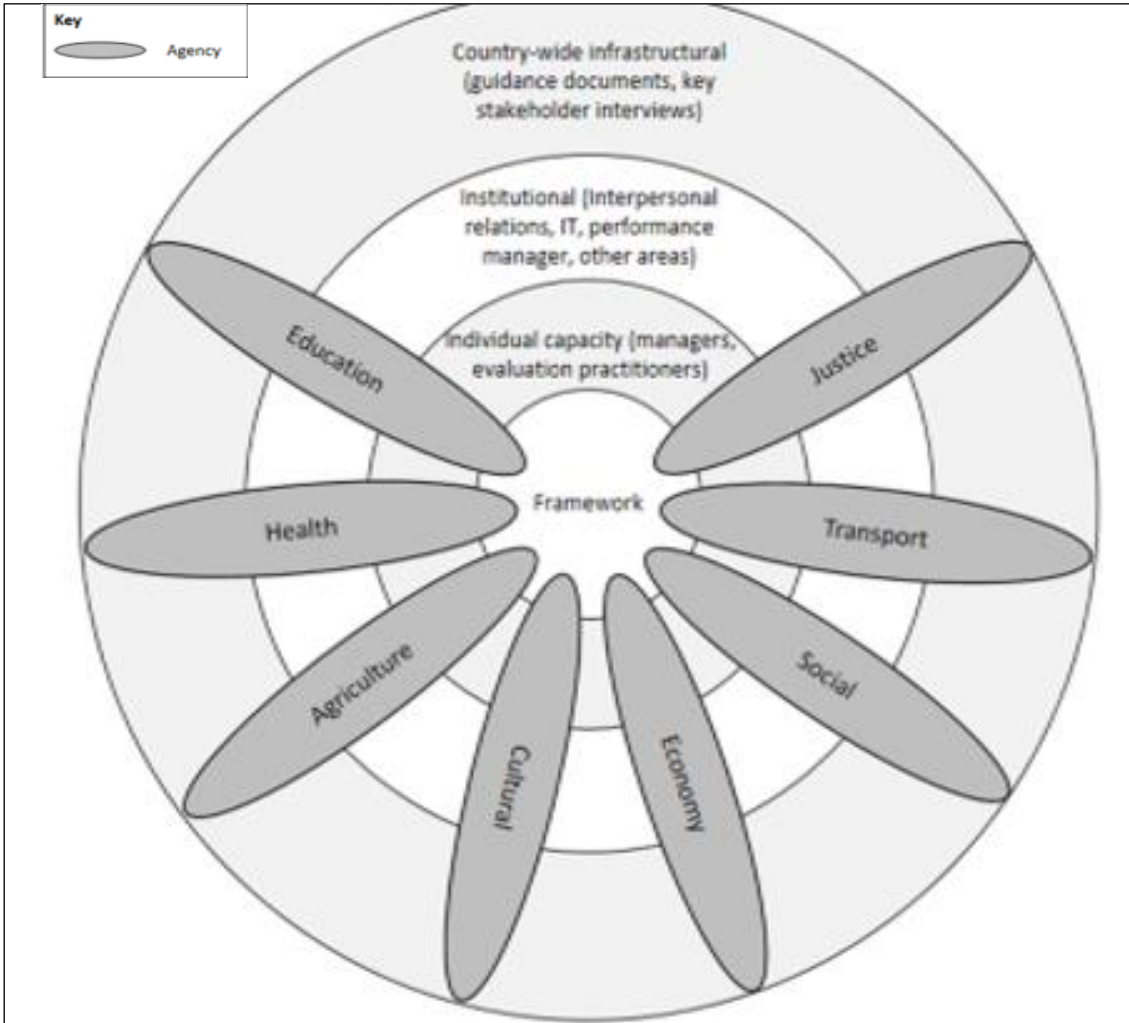
The case study object of the research is the architecture and use of strategic results frameworks cascading from country-level guidance to use by actors (managers and performance/evaluation specialists) in sectors or agencies within each country case study setting and the associated management and governance systems. The experiences and events within the contextual layers (that is, *ideas, interventions and frameworks; institution* (line agencies); and *wider infrastructure* settings) were identified, and the underlying mechanisms analysed. The capacity of managers and evaluation practitioners within the contextual layers were examined, and the nature and role of development, management, evaluation and governance considered. Conclusions for development effectiveness and governance were drawn out from the case studies.

This approach to case study research was referred to as theory-testing by George and Bennett (2005), critical by Yin (2009) and theory-seeking by Bassey (1999). This research encompassed a combination of theory-testing and theory building seeking to extend the knowledge in the architecture and use of strategic results frameworks in different settings as the literature review had demonstrated a knowledge gap in this area.

Within the scoping phase, my research showed that there were more considerations than documented in the literature. Hence as the research progressed, the shift to theory-building became more apparent.

**3.4.4. Case study data collection**

Figure 9 displays the structure of the A/NZ case study object. This structure of the case study research was applied to the case study research undertaken in two countries (A/NZ and PNG). The case study structure had to be adapted in PNG due to the country-level approach taken for the PNG country and sector results framework. Prior to the case study research, country-specific performance management, evaluation, budget cycle and other relevant documents were scanned to provide background information.



**Figure 9. Structure of country case study object for A/NZ case study research.**

The structure of the country system research comprised of guidance and interviews from central government agencies and donors, which operate in the outer *wider infrastructure layer*). Eight key sectors or agencies are the implementing institutions (the research included a document review of each sector/agency SOI, their and their strategic-level results framework, and annual reports). Interviews were t undertaken in both PNG and A/NZ agencies. As part of the deeper case study in A/NZ agencies, ratings

of key concepts relevant to the architecture and use of strategic results frameworks were completed by managers and evaluation practitioners.

For each case study, research included the following sectors/agencies: health, education, justice, transport, social development, economic development, agriculture, and up to two other agencies (those with a cultural focus). A rationale for selecting these agencies was that during the scoping phase, key informants recommended me to select the same agencies or equivalent for each case study to provide a systematic basis for the research and analysis processes.

For each sector/agency, the following types of data were collected:

1. Document review: SOI, annual report and strategic plan. The criteria were identified from the literature and practice. A rating matrix was used by the researcher to rate each agency's SOI and annual report against criteria which drew from information from the literature review on the architecture and use of results frameworks (see Appendix D for data collection tools accompanying rating matrix and Appendix B for the formative diagram underpinning the document review matrix).
2. Qualitative data – face-to-face interviews with national agencies included Treasury, SSC (or equivalent), Department of Planning, M&E (or equivalent), key government departments (health, education, social development, economic development, transport, justice, agriculture, etc.), donors. Two face-to-face interviews (individual or group) with personnel from two areas of each sector/agency included: (i) management and (ii) evaluation. Interviews had three parts, incorporating aspects of grounded theory, including:
  - i. Open-ended questions on the architecture and use of frameworks management, and evaluation approaches
  - ii. Impact model result rating table (A/NZ case study only)
  - iii. Further open-ended questions on the impacts, significance and the implications for practitioners, and governance.

The links between qualitative findings in the interviews and results ratings were examined iteratively within a mixed-method approach (Tashakkori & Teddlie, 2003)

The case study interview guide differed from the key informant interview guide used in the scoping phase (Appendix D: Interview guide 1). The initial key informant interview guide was derived directly from the impact model (drawn from literature) on the intended use of results frameworks. Given the emergent findings of mixed knowledge and



understanding by key informants in the initial part of the scoping phase, the interview guide was updated (Appendix D: Interview guide 2) with the incorporation of more open questions on development, evaluation, M&E and the impacts of the changes for management, governance and evaluation.

### 3.5 Data analysis

Within each case study: for up to eight agencies: data analyses included:

- i. A content analysis approach (Miles & Huberman, 1994) was used to analyse the annual reports and SOIs, using a good practice criteria matrix drawn from the literature review.
- ii. The SOI and the annual report was analysed against good practice criteria using document review matrix drawn from the literature review. The first level of data analysis in each case study included collating results from the document review of agency documents and rating survey. Pattern matching (Miles & Hubermann, 1994) was undertaken through triangulation between the document review, interview findings, and rating table (New Zealand) for each case study.
- iii. The impact model provided a structure for incorporating elements of a deductive approach. This research encompassed a table of the intended key results (refer Appendix D) and a rating scale. After the first part of the interview, stakeholders were asked to rate progress towards these results in their agency.
- iv. Interviews were held with managers and evaluation personnel of each agency using an inductive approach and aspects of grounded theory. This analysis involved the secondary coding of the qualitative findings within each case study. The open categories were further analysed using axial coding (Straus & Corbin, 1998).

The first part of the analysis process included the coding of the qualitative findings (in NVivo software) using the emergent analysis frames incorporating an adapted use of Pawson's (2006) contextual layers as a new typology of (1) *ideas, interventions and frameworks*; (2) *institution* (from line agencies); and (3) *wider infrastructure* settings (including governments' central agencies). *Interpersonal relations and individual capacities* concepts were embedded within each of the three layers of the analysis frame, as during my scoping phase analysis, I noted these concepts were findings within each layer. I then identified thematically key conceptual areas to use as code frames within

each of the reality layers (Figure 10). The data analysis and initial thesis write up were completed between 2015 and 2018.

Name	Sources	References
Infrastructural - political support, funding, resources	1	5
Accountability - performance management, audit, evaluation, financial, country systems	20	168
Central agencies	14	60
Development - approaches, power, culture, national capacity	8	32
Donors - govt to govt cooperation, role of monitoring & evaluation	5	34
Governance - national, donors, evaluation	8	38
Institutional	1	4
Individual capacity	1	3
Evaluation practitioners capacity	8	23
Other agency personnel capacity-capability	16	45
Interpersonal relationships	1	3
Communication, links, management administration, roles, IT systems	17	74
Management- strategy leadership, policy, learning, decision-making, culture, resources, beneficiaries	25	310
Role of Ministers	9	21
Monitoring and Evaluation, research within institution	24	223
Performance management within institutions, audit, aggregation, reporting	25	287
Value for money	16	41
Interventions-frameworks-ideas	1	3
Architecture - Structure of frameworks	22	173
Ethnic subgroups disadvantaged other subgroups	14	66
Evaluation role and nature	13	70
Use	19	102
Measurement	18	82

**Figure 10. Emergent code frames in NVivo.**

### 3.5.1. Mixed-method analysis process

The mixed-method analysis of the findings was undertaken in three steps.

First, the ratings were collated from agency personnel and document review. Qualitative data was coded under the emergent codes identified in scoping phase and A/NZ case study. The PNG interviews were coded using the emergent code frames displayed in Figure 10 with additional codes added such as country planning. The second case study for PNG was undertaken at a lighter level due to availability, access and scope of the research.

The second level of my analysis used the critical realist stratification by Wuisman (2005) of empirical (experiences), the actual (events) and the real (mechanisms) within the three layers. Burgess (1984) outlined that by analysing nested elements, a theory can be built within case studies. The coded data was analysed for experiences (descriptive), and the associated events and underlying mechanisms (interpretive and evaluative) within the contextual levels of *ideas*, *interventions* and *frameworks*, *institutions* and *wider infrastructure* layers. I decided to keep the experience and event findings together as the data became too segmented. The secondary coding of the qualitative findings under the code frames was undertaken using a matrix table format displayed in Appendix E.

Third, a mixed-method analysis of the document review, rating table and qualitative data was undertaken by triangulating the multiple data source findings to draw out concept themes. Further analysis within the three social reality layers examined links between the experiences (descriptive) and the impacts (interpretive) and the significance of the events, mechanisms and implications (evaluative) for development, management and evaluation. This analysis provided more in-depth explanations as well as identifying the unintended findings.

Induction, deduction and abduction (Figure 8) were deployed as analysis techniques in this research as promoted by Wuisman (2005) to examine the links and relationships between and within the social reality layers. I used induction and deduction initially to identify the experiences (descriptive) and draw out events that appeared causally linked to the experiences. Abduction was used to analyse the relationships between experiences and events and events and the mechanisms within a country system. A sub-theme analysis was undertaken within each emergent code in A/NZ for consideration of Māori and Pacific peoples in development, management and evaluation.

A comparative pattern-matching analysis of findings (document, interview findings and rating table) was undertaken across the case studies. The data from multiple case studies using nested contextual layers (*ideas, interventions and frameworks, institutions and wider infrastructure*) stratified by the critical realism strata (that is, experiences, events and mechanisms) were used to answer the research questions and contribute to the body of knowledge in this area.

A cross-case analysis was then undertaken based on findings in each case study with the key question of 'what works for whom, where and why' (Pawson, 2006) in different contexts. The validity of and confidence in the findings were based on triangulation of multiple data sources, multiple methods, and from multiple perspectives.

### **3.6 Research method limitations**

There are some caveats on these findings. First, the review of accountability documents was undertaken as part of the A/NZ case study and included both central and agency documents. The document review was undertaken at a lighter level for PNG as accountability documents, other than at the centralised level, were not so readily available.

My research was initially designed from my experience as an evaluation professional with a background in academic studies. It was premised on expectant knowledge and practice that I considered might be forthcoming from documents and the research participants. However, the key informant interviews and background documentation in the scoping phase made it clear that knowledge and practice concepts were more

complex than indicated in the literature and expected from practitioners. I also considered that maybe I was placing too much researcher bias on the research design,

Besides, I now consider that my earlier intellectual awareness and thinking may have constrained how I initially approached my topic. At the beginning, I was relatively unaware of underpinning theory and literature despite my supervisors trying to provide constructive feedback. This reflection also contributed to my realisation that, as a practitioner, I had used knowledge and practice concepts unreflexively, without deeper intellectual understanding.

As a result of the various reflections and reconsiderations during the scoping phase, I shifted my focus to incorporate aspects of grounded theory with more open-ended questions. This amended interview guide assisted in mitigating some researcher bias in the case study phase. My amended approach appeared to support participants to share existing and emergent concepts which, in turn, assisted me to put forward existing and emergent concepts.

There may have been limitations on my analysis of the PNG case study as it was undertaken in English with a Tok Pidgin speaker present. PNG interviews in Tok Pidgin were translated at the time of the interview, and the interview transcripts transcribed in English. The data collection tools were translated into Tok Pidgin for the PNG fieldwork. In addition, story-based approaches such as *talanoa* (Farrelly & Nabobo-Babat, 2014) for qualitative interviewing were incorporated into the case study research in both PNG and A/NZ. This approach was recommended to me by a Samoan researcher as it resonates for Pacific people given their preference for verbal communication.

One acknowledgement for this research included that I was working in evaluation and management in both PNG and A/NZ contexts (2010–2016) during the time this research. I endeavoured to keep my research separate from my work inputs. I consider there were no conflicts of interest in undertaking my fieldwork. However, my understanding of case study contexts was augmented from my work, particularly in PNG and about sustainable development as I became involved in contributing to the MTDP 2 released in 2015 after my fieldwork was completed. This latter work gave me some insights into the practice contexts for these research findings.

There were some challenges encountered in PNG with agency staff being available. Hence the lighter case study research was undertaken. However, I am very appreciative of the time and inputs from PNG research participants as I did not want this research to be intrusive or extractive. A research summary will be shared with all people who participated in this research.

One limitation noted during my analysis was that policy specialists were not included in the research participant sample. Their inclusion may have given more input on how development theory and thinking was developed and applied.

Regarding the literature and documents reviewed, there were time limitations on how many documents were sourced and their availability. In addition, given the scope of this topic was significant, no community or regional level research was undertaken, which may have provided broader perspectives and further insights.

### **3.7 Overview of the evolution of the research**

This overview of my research journey highlighted the critical steps where the research direction was clarified.

1. Initial interest was displayed in both the value of strategic architecture in the governance of development, and the idea of developing good practice guidelines, and the extent to which these might be different or similar across different jurisdictions
2. leading to a search for a theoretical frame against which to interrogate various strategic architectures and identifying Brinkerhoff's (2006) impact model as a starting point in relation to the OECD-DAC Aid and Development Effectiveness Principles (2005) the intended impacts and then considering
3. the use of critical realism (Bhaskar, 1998) as the theoretical perspective aligned with Wuisman's (2005) idea of the "stratification of reality" (all of which was broadly discussed in Chapter One).
  - i. This was followed by a more detailed literature review undertaken in three parts that included:
  - ii. the identification of key concepts from development, management and evaluation knowledge fields (presented in Chapter Two)
  - iii. a review and synthesis of development, management and evaluation approaches in national development contexts in two selected countries recognised in the literature review as demonstrating good practice. (presented in Chapter Four)
  - iv. a document review on the practice context of A/NZ and PNG country approaches to development, management and evaluation (presented in Chapter Four).
4. Once these initial steps had been completed, the primary research was undertaken (and the methodology for this is described in Chapter Three). It comprised two distinct phases.

- i. A scoping phase that involved 18 key informant interviews in gaining an understanding of context, perspectives and existing knowledge and practice. This resulted in an emergent research design using grounded theory, as there were more unanticipated findings beyond the dimensions included in the impact model, which was updated for the case study research undertaken Phase Two.
  - ii. The case study research in A/NZ and PNG comprised Phase Two and included qualitative interviews (36) with managers and evaluation practitioners and funders and the rating of impact model dimensions that was completed by A/NZ case study research participants. The interview guide used for case study interviews incorporated more open-ended questions to draw out the extended concepts from the research participants. Deeper research was undertaken in A/NZ with more interviews and more in-depth review of the accountability documents. The lighter level was due to more limited availability of research participants and line agency accountability documents.
5. Following the data collection, the analysis processes included the coding of the qualitative findings (in NVivo software) using an emergent analysis frame which incorporated a new typology which was initially adapted from Pawson's contextual layers (2006) of (1) *ideas, interventions and frameworks*; (2) *institution*; and (3) *wider infrastructure* settings. Interpersonal relations and individual capacity settings were embedded within each of the three layers of the analysis frame. I then applied this new conceptual research frame as a new typology for this theory building research on national planning, management and evaluation.
6. At this point, the description of the substantive and emergent findings was undertaken (Chapters Five, Six and Seven) under the three contextual layers and the knowledge concept areas.
7. Then followed a critical and reflective discussion of findings, and the development of the idea that an integrated epistemology for development, management and evaluation could extend knowledge and enhance practice (Chapter Eight).
8. A discussion of the development of an idea for emergent new paradigm integrating management and evaluation of 'evaluative management' and an integrated evaluative management model (Chapter Nine).
9. Conclusions and implications from the research were considered including areas for further research (Chapter Ten).

### **3.8 Overview of this research**

After an initial scoping phase, an evaluative case study approach was used as a method in the context of a critical realist perspective that was abductive and emergent. The research phases included:

1. **Phase 1 – Scoping: Key informant interviews** (18) from central and line agencies, wider stakeholders; and a literature review.
2. **Phase 2 – Comparative analysis – document review and two case studies** (included 36 interviews).
  - iii. A/NZ – deeper case study – document review of accountability documents, interviews (19) managers (12), evaluation specialists (7), and rating table of main components by interviewees (12). The research scope included eight separate line agencies.
  - iv. PNG lighter case study – Document review, interviews (17) from central and five separate line agencies.

In Chapter Four, the practice contexts of the two case studies were outlined.

## Chapter 4: Practice contexts

This chapter highlighted the practice contexts of the two case studies (A/NZ and PNG). The practice contexts for this research included national governments' development planning, management and governance, MIE requirements, and the interface with development partners and donors from which nations receive development support. A sub-theme of this research was how national governments approached development for indigenous subgroups within national development contexts such as Māori and Pacific peoples in A/NZ and provinces (cognisant of tribes) in PNG. Understanding the evolution of thinking and practices underpinning each country approach provided context for this research.

### 4.1 Aotearoa/New Zealand country governance and management

For understanding the current governance arrangements in A/NZ, a highly simplified version of this history must suffice. The short-hand history suggested that, before colonisation, the structure of early Māori society was through people identifying as iwi (tribes) with the waka (canoe) they arrived on (Te Ara: The Encyclopaedia of New Zealand [Te Ara], 2006). Sub-tribes (hapū) formed as growth occurred, and iwi and hapū spread out to different regions. From the arrival of Europeans in the 1800s, Māori social organisation changed again with trade and arrival of the missionaries. The attempt, in 1835, to develop an independent relationship with Great Britain resulted in *He Whakaputanga o te Rangatiratanga o Nu Tirene*, the Declaration of Independence of the United Tribes of New Zealand. Although accepted by the Crown, it was superseded in 1840 by the Treaty of Waitangi (the Treaty). After the Treaty was signed, constitutional changes meant that in 1841, New Zealand became a crown colony separate from New South Wales (Ministry for Culture and Heritage, 2017). However, the Treaty had two versions that differed in meaning between the English and Māori texts, which led to the loss of Māori sovereignty. The colonial A/NZ Government did not uphold the Treaty, and a court-case judgement in 1877 declared the Treaty to be a "simple nullity" (Ministry of Culture and Heritage, 2019). Between 1880 and 1900, the Māori population declined through "destruction, disease and despair" (Te Ara, 2006, p. 76) with land wars occurring from 1845 to 1872, and further land confiscations. By the second half of the 19th century, Māori had lost control of much the land they had owned.

From the 1950s, Māori increasingly placed pressure on the A/NZ government for redress for land confiscations and increased rights to sovereignty. The Treaty of Waitangi Act passed in 1975, and the Treaty of Waitangi Tribunal established which started researched breaches of the Treaty. From 1980 onwards, this tribunal has led to iwi and



hapū reaching settlements with A/NZ government as part of redress for past grievances by the Crown.

In 1987, the A/NZ Māori Council (A/NZ Māori Council v Attorney-General) High Court case occurred where the Māori Council sought enforcement of section 9 of the State Owned Enterprises Act 1986, which read: “Nothing in this Act shall permit the Crown to act in a manner that is inconsistent with the principles of the Treaty of Waitangi” (State Services Commission, 1988, p.14). This case then went to the Court of Appeal, and a judgement upheld the Treaty principles (Haywood, 2012). In 1989, the principles were set out by the Labour Government, which included:

- i. the government has the right to govern and make laws
- ii. iwi have the right to organise as iwi, and, under the law, to control their resources as their own
- iii. all New Zealanders are equal before the law
- iv. both the government and iwi are obliged to accord each other reasonable cooperation on major issues of common concern
- v. the government is responsible for providing effective processes for the resolution of grievances in the expectation that reconciliation can occur. (p. 6)

Wheen and Haywood (2012) noted that “by 2011 no later government had defined any new treaty principles, although some (such as the National government in 1991) have reflected on the 1989 principles” (p. 19). Subsequently, other legislation refers to the principles of the Treaty translated into social policy (such as Education Act, 1989) and environmental policy (such as Conservation Act in 1987). The Treaty principles evolved through the Waitangi Tribunal to the Court of Appeal and into government in A/NZ through case law.

From 2008 to 2015, there was a strong movement among Māori tribal and non-tribal organisations for constitutional change that reflects the Treaty of Waitangi, which was initiated by Māori leaders, who facilitated 252 hui (discussion sessions) nation-wide from 2008. An independent working group proposed engaging with the Crown and local government over the need for constitutional transformation and a Te Tiriti (Treaty) convention by 2021 (Tawhai & Gray-Sharp, 2011; *Matike Mai Aotearoa – The Independent Working Group on Constitutional Transformation*, 2016). Among the panel’s recommendations and of relevance to this discussion was that the government affirmed the Treaty as the foundational document. New Zealanders were invited to continue the conversation about the place of the Treaty in the A/NZ constitution; investigate how Māori parliamentary representation may be improved and how local government processes

can better reflect the views of *tangata whenua* (indigenous peoples). However, no timeline was proposed to progress this constitutional conversation.

#### **4.1.1. State sector management**

Background to the changes undertaken in the 1980s was A/NZ's history of being viewed as a welfare state as in the 1960s the government post-war used an interventionist type approach for the role of government and subsidies to protect A/NZ's economy. An underpinning ideology was that A/NZ citizens had rights to access services and maintain a quality of life (Dalziel & Saunders, 2014). However, the economy struggled through the 1970s and into the 1980s with "progressive deterioration" (Schick, 1996, p. 13) of the economy occurring

From the mid to late 1980s, A/NZ underwent a period of change in the state sector driven by a desire to reshape the economy, societal structures and A/NZ position internationally (James, 1992). A key driver for change in the public management system was economic imperatives. Other drivers for change were a shift in perspectives about the role of Government in the economy, and a desire to make the state sector more responsive to ministerial demands (Schick, 1996). Scott (2001) outlined that this new model was based on private sector approaches using management authorities and board accountabilities, to streamline processes. This involved government reforms such as deregulation and a shift to a free market with removal of farming subsidies and privatisation of state assets which impacted significantly on A/NZ businesses used to a more government subsidised economy. The changes in government leadership contributed to these extensive changes being undertaken in A/NZ by a Labour government which were regarded internationally as significant (Schick, 1998) at the time.

From 1990, development emphasis in A/NZ shifted towards economic growth and efficiency in resource allocation and results. This shift included devolving managerial power to agencies which were an influencing factor in A/NZ's development and management processes. Cook (2004) outlined that the theoretical basis for the changes in A/NZ, based on economics as the reforms identified challenges that to address in order for public management systems to become more efficient and effective. Underpinning the responses by A/NZ Treasury were five key principles promoted in *Government Management* (The Treasury, 1987).

The five core principles for A/NZ's public management system (Cook, 2004) included:

1. **Clarity of objectives** – Managers were expected to specify the objectives, the process to achieve these objectives and the managers who are responsible.
2. **Freedom to manage** – Managers were to be given the freedom to allocate resources to achieve these objectives most efficiently.

3. **Accountability** – Managers must be accountable for the decisions they make, and those on whose behalf they act must have the means to make that accountability ‘stick’.
4. **Effective assessment of performance** – Managers needed to assess progress towards objectives, establish the quality of the performance, and identify the external factors over which they have no control.
5. **Adequate information flows** – there needed to sufficient information on performance to assess the quality of resource decisions. (p. 4)

These operationally focused principles (SSC, 1988) set the context for devolving public management authority to A/NZ agencies and the expectation of effective assessment of performance using adequate information to assess the quality of resource decisions. This research examined the impact and significance of these principles at A/NZ central government and agency levels in comparison to a more centralised approach used in PNG.

Figure 11 (Gill, 2008, p.4) displayed the 1998 reformed management accountability relationships in A/NZ under which the state sector operated. According to Treasury (Cook, 2004), differences in this model were that Chief Executives and department were to be held accountable for the achievement of their objectives and associated structural changes required.

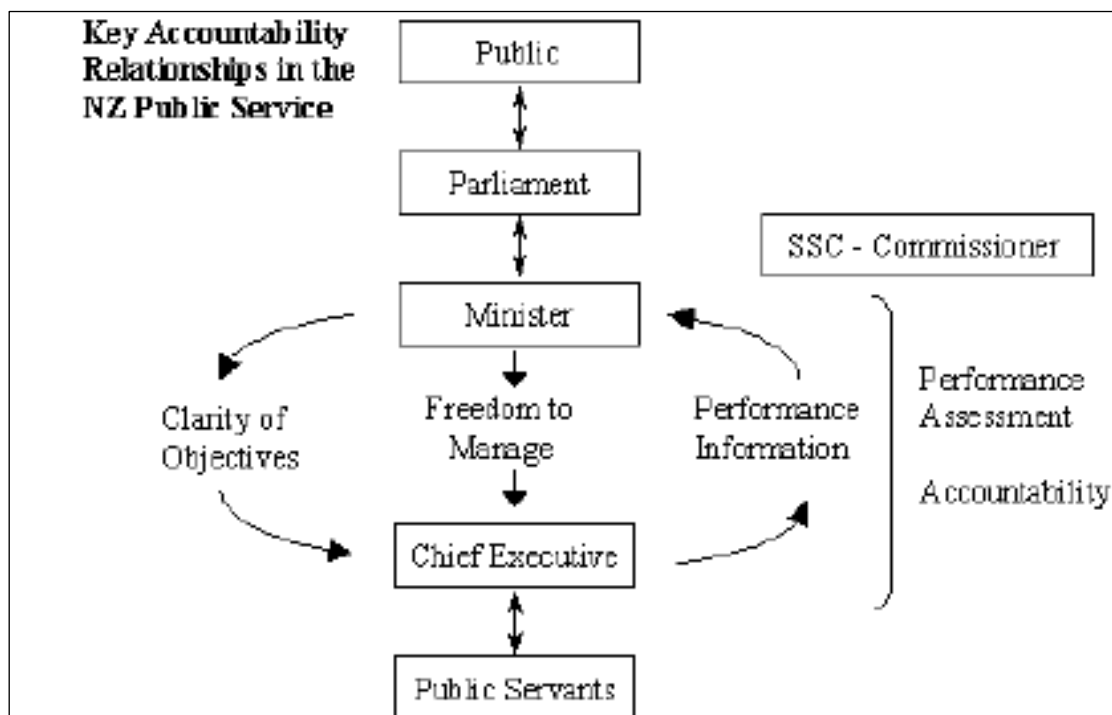


Figure 11. Key accountability relationships in the A/NZ Public Service (Gill, 2008, p. 40).

In the 1990s, the performance focus moved higher to Strategic Results Areas (State Services Commission, 1998a) under the National Party. This shift appeared influenced by a growing international focus on managing for outcomes (Canadian Centre for Management Development, 2001). However, it was discarded by the incoming Labour-led coalition (Ryan, 2003). The reluctance of the government to explicitly link relationships between strategy and outcomes were noted in the *Review of the Centre* (Advisory Group, 2001), and an increase in strategic conversations recommended. Alongside these changes, Ryan (2003) also noted that shortcomings in evaluation in A/NZ were also becoming evident.

However, by 2004, Treasury (Cook, 2004) had already noted that these changes led to the separation of policy from operations. Ryan (2003) considered the A/NZ 1980s reforms focused on “economic rather than managerial effectiveness” (p. 7) as outputs were the basis of accountability for budgetary and financial management. These outputs were purchased in contracts from agencies and chief executives by ministers and funders. Cook (2004) considered these approaches led to a focus on processes and outputs as measures of performance (Boston & Pallor, 1997).

#### **4.1.2. Managing for Outcomes**

To address these growing concerns, the A/NZ State Services Commission (SSC) from 2001 to 2003 piloted the *Managing for Outcomes* initiative (SSC, 2003). Since 2001, expectations that *Managing for Outcomes* would be implemented in public service departments. By 2004, all public service departments and Crown entities were mandated to manage for outcomes, and the guidance outlined that agencies were to use of M&E information within their management approaches (SSC, 2003). The Public Finance Act 1989 (amended in 2004) and Crown Entities Act 2004 reinforced these expectations as the language of the two Acts were extended to include outcomes, impacts and objectives, and were referred to by the term “results”. A results focus was extended across the A/NZ state sector, including government and Crown agencies. However, in my review of the A/NZ evaluation and management guidance before 2000, it appeared there were limited references made to the Treaty of Waitangi and outcomes for Māori.

#### **4.1.3. Outcomes for Māori**

From 2002, evaluation in A/NZ was increasingly being called upon to consider the cultural appropriateness of services provided for Māori (Thomas, 2002). Te Puni Kōkiri (TPK), the Māori Development agency produced a set of guidelines (TPK, 1999) on evaluation for Māori. The guidelines described a foundation for evaluating performance for Māori, based on the Treaty, the government’s strategic priorities, and the priority accorded to Māori as a target group. Cunningham (2003) had emphasised that

“responsiveness, in terms of Māori knowledge, needs to be operationalised if it is to contribute to Māori development (p. 63).

Durie (2004) argued that:

Ultimately the impacts of public sector reforms on Māori must be measured against both sets of criteria. First, as citizens of New Zealand, Māori performance should be assessed according to universal outcomes (such as life expectancy, educational achievement, employment). There should not be wide disparities between groups. This aim is not necessarily a consequence of the Treaty of Waitangi but of the goals of a fair and just society.

Second, as an indigenous people Māori performance should be measured against Māori specific outcome indicators that are derived from Māori culture and traditional physical resources such as land. Best outcomes for Māori are the product of universal and Māori specific outcomes. (p.15)

In response to a sense that universal indicators alone might not be serving Māori well, Kingi and Durie (2000) developed a culturally appropriate assessment tool for Māori Mental Health outcomes using a *te Whare Tapa Whā* (a four-walled house) model and developed specific indicators for four cornerstone dimensions (physical, spiritual, family and mental health) for use in mental health interventions. In 2001, TPK developed a monitoring framework that included indicators and four levels of application (local, regional, national, and international) for Māori development. Durie, Fitzgerald, Kingi, McKinley & Stevenson (2003) emphasised that assessing Māori development needs to include both generic and specific approaches.

In 2006, Cram noted the growth of awareness on impacts of colonisation and the growing voice of Māori. Cram, Pipi and Paipa (2018) outlined the journey of the growth of Kaupapa Māori Evaluation in A/NZ incorporating an epistemology of undertaking evaluation a Māori way. Carlson et al. (2017) in their research provided contextual examples of the utilisation of *Kaupapa Māori Evaluation* [KME] and detailed the importance of “mutually beneficial relationships, power-sharing and collaboration” and “how these play out in the real world of KME” (p. 75).

As a Pākehā (European) researcher, it was important to me, given the historical significance of the Treaty of Waitangi, the bi-cultural, and multi-cultural heritage of A/NZ, that cultural approaches and appropriateness of A/NZ’s approach to development, management and evaluation were examined as part of this research.

#### **4.1.4. Outcomes for Pacific peoples**

In A/NZ, there are 22 Pacific nations represented, each of which has differing cultures, languages and beliefs. In 2014, Pacific peoples made up over seven percent (295,941) of the total population (Statistics NZ, 2014). From the 1950s to 1980s, there were waves of Pacific peoples recruited to A/NZ for work. Further immigration occurred as workers settled and were joined by family members. A/NZ has had a mixed history with Pacific migrants. In the 1970s immigration policies resulted in 'overstayers' deported despite incentives for Pacific peoples to come to do work that was not popular with existing residents: young men were recruited for hard agricultural and forestry work, young women as domestics, and many more into manufacturing before the oil crisis and A/NZ's exclusion from the EU. Pacific peoples in A/NZ have had to strive to gain a voice on matters relevant to their communities which included successfully challenging immigration policies at the Privy Council in London (Ministry for Pacific Peoples, 2017, p. 8). Since 1980, the A/NZ government established entities for Pacific people. Pacific people have become more embedded in A/NZ society in sport, public sector and politics.

As part of this growing awareness of Pacific perspectives, a model was created by Pulotu-Endemann (2001) for identifying specific outcomes for Pacific peoples. This *fonofale* model incorporated values and beliefs from Pacific cultures. The four posts in the *fale* (Samoan word for house or building) represented different dimensions including spiritual, physical, mental and other (such as gender). This model promoted Pacific ways (family, culture and spirituality) valued by Pacific peoples for service delivery and included environment, time and context dimensions.

However, in a study on inequality for both Pacific and Māori (Marriott & Sim, 2014), indicators showed that while there is a decreasing gap between Māori and Europeans in some areas, the indicators for Pacific peoples show the inequality gap (for example – life expectancy) with Europeans has increased over the past ten years. The gap has also increased for Māori in some areas, such as secondary school retention and youth unemployment. My research included a sub-theme examining how considered are Māori and Pacific peoples in A/NZ development, management and evaluation approaches.

## **4.2 Performance management and evaluation**

Since the 1990s, A/NZ public services agencies were placed under increasing public scrutiny. As part of the *Managing for Outcomes* mandate, agencies were requested to use an intervention logic approach (Ryan, 2002) in their statements of intent (SOIs) as part of budget processes. This approach was intended to shift the focus from outputs to include impacts and outcomes. These developments appeared to be influenced by results-based management approaches being used internationally (Rist, 2004) and by

Australasian research and evaluation specialists such as Duignan (2007) with his use of outcome frameworks in A/NZ, and Rogers (2008) in Australia on programme theory for evaluating programmes and projects.

From the 1980s to 2010, emphasis on performance and evaluation varied (Duignan, 2002). Lunt, Davidson & McKegg (2003) emphasised the value from using more systematic evaluative approaches and advocated of a formalised role for evaluation by the A/NZ government. Ryan (2003, 2011) noted that some A/NZ government agencies had a historic interest in outcomes and evaluation including the Ministry of Social Development, Department of Labour and Department of Social Welfare, which have “albeit limited programmes of evaluation and did publicize results” (2003 p. 8).

Ryan (2011) outlined that “most government spending is on programmes with the proportion spent on public service operations being relatively small. The greatest potential for savings lies in cutting programmes, but this can be electorally unpopular” (p. 22). Furthermore, he emphasised that central to prioritisation of programmes, “good and extensive evaluation of the programmes under consideration” was required. However, “not many governments have such information” (OECD, cited in Ryan, 2011, p. 23), Ryan also highlighted that this lack of evidenced-based information was an issue in A/NZ “with so little evaluation conducted” (p. 23).

Funding to A/NZ agencies occurs under a ‘vote’ system where each agency is allocated several ‘votes’ comprising of a group of output classes for which funding is allocated for by Treasury. This approach means that each agency implements the budget allocation within its agency intentions. However, Hughes and Smart (2012) considered “this requires a change in the way success is understood at present. The outcomes currently presented in agency Votes are often lofty and seldom achievable” (p. 7).

Hughes and Smart (2012) noted that “over the past 30 years New Zealand’s system of public management had seen a number of positive changes, both systematic and incremental” (p. 203). Yet, they noted that “previous attempts at establishing an outcomes focus, such as *Managing for Outcomes*, were not sustained because of the focus on producing outputs. They considered that:

the new ways of tackling performance management will require the public sector to develop strategy across sector groups, describe outcomes so they can be reliably measured, improve the use of evaluation as a tool to inform performance frameworks, and measure success. (p. 7)

Furthermore, Nunns, Peace, and Witten (2015) in their review of publicly accessible evaluation reports highlighted that around half of the 28 evaluation reports analysed “lack an argument” (p. 159) based on the evidence which can impact on the validity and quality

of evaluative judgements and conclusions. These findings, which they considered required further exploration and international comparison, raised questions over the validity of evaluation reporting in A/NZ.

In 2017, the A/NZ government produced a handbook on evaluation (Superu, 2017) “designed for people funding or purchasing social services, and providers of social services in New Zealand who need to understand more about how to measure and understand the effects of their programmes or initiatives” (p.1). However, this unit was closed in 2018. As of 2019, the role and nature of evaluation in A/NZ remained unspecified, and there was no mandated approach for evaluation by central agencies.

#### **4.2.1. Aotearoa/New Zealand case study research**

A/NZ’s public management system contained five core principles namely “clarity of objectives, freedom to manage, accountability, effective assessment of performance and adequate information flows” (Cook, 2004, p. 4). The State Sector Act (1989) outlined that accountability and performance are devolved to each agency chief executive. However, from my literature review, I considered that these core principles were more focused on management processes rather than higher-level guiding principles which if made explicit could enhance inclusion of different stakeholder perspectives and needs.

In A/NZ, the States Services Commission is the central agency responsible for issuing accountability and performance reporting guidance in conjunction with The Treasury and Department of Prime Minister and Cabinet (DPMC). The Treasury is responsible for the budget allocation, and individual agencies provide an account of funds expenditure in their annual report. Audit NZ is responsible for auditing financial information and, since 2012, has extended its functions to auditing for non-financial information.

However, no central agency is responsible for M&E under the managerial devolution which occurred in the 1980s reforms resulting in the State Sector Act 1989 and Public Finance Act 1989. There was also further devolution by agencies for service provision to Crown entities which was outside of the scope of this research.

Central to my research was an examination of the nature and role of evaluation from central agency to implementation by eight main public sector agencies. I focused on key government agencies of health, education, agriculture, economic development, transport, justice, Māori development and Pacific Island Affairs where the latter two agencies are mandated with cross-agency responsibilities for the two key population groups.



## 4.3 Papua New Guinea

PNG was selected purposefully (Miles & Huberman, 1994) as a developing country case study as it utilised a national-led approach to country development. PNG also had agreed to incorporate OECD (2005) aid effectiveness principles in their national development planning approach. A fundamental aspect of the aid and development effectiveness principles was that developing countries would develop their own national strategic results frameworks and that donors would align to the country led development. This partnership alignment was examined in my research.

PNG, with its diverse geography and “highly culturally and linguistically divided society” (Chand, 2010, p. 3), has a history of centralised development planning with a national planning approach focused on supporting national cohesiveness and resource allocation alongside service delivery expansion across provinces. PNG continued to work on improving literacy rates, maternal and infant health, and life expectancy, but experiences challenges from its reliance on resource extraction and accountability issues in ensuring funds reached the provinces for essential services such as schools and hospitals (UNDP, 2014).

### 4.3.1. Background

PNG gained independence from Australia in 1975 after movements which were active in the lead-up to independence and formation of the nation” (Chand, 2010) although none of the movements were particularly aggressive. The PNG Constitution (PNG Government, 1975) contains the following underlying “National Goals and Directive Principles” (p. 2):

1. **Integral human development.** We declare our first goal to be for every person to be dynamically involved in the process of freeing himself or herself from every form of domination or oppression so that each man or woman will have the opportunity to develop as a whole person in relationship with others.
2. **Equality and participation.** We declare our second goal to be for all citizens to have an equal opportunity to participate in, and benefit from, the development of our country.
3. **National sovereignty and self-reliance.** We declare our third goal to be for PNG to be politically and economically independent, and our economy basically self-reliant.
4. **Natural resources and environment.** We declare our fourth goal to be for PNG’s natural resources and environment to be conserved and used for the collective benefit of us all and be replenished for the benefit of future generations.

5. **PNG ways.** We declare our fifth goal to be to achieve development primarily through the use of PNG forms of social, political and economic organization. (p. 2)

Chand (2010) outlined that cultural diversity and the location of resources led to democratic and political instability. As after independence, successive governments did not last a full term in office until 2002 when Somare was elected as Prime Minister.

#### **4.3.2. Development planning and management**

The Central Planning Office was established in 1963 by the colonial administration of the then Australian Territory of Papua and New Guinea (Project Planning Team, 1963). Planning gained increasing economic prominence under the guidance of an economic adviser where statistics were gathered from economic sectors and based on that, the first project plans (1963–1967) pre-independence were drawn up. As part of the shift to independence and given the tribal nature of PNG with over 600 tribes, a nationally led planning approach was selected as a tool to support development planning and resource allocation. This approach has underpinned successive development plans and associated processes (National Research Institute, 2008).

The PNG Government (GoPNG) undertakes its centralised development planning through its Department of National Planning and Monitoring (DNPM) with devolved governance and management to 22 provinces, 319 districts and over 600 local level governments, recognised under Organic Law passed in 1998.

However, in 1999 Igara, Secretary of the Prime Minister's Department under Prime Minister Moratua highlighted that:

Reforms were clearly needed. It was a devolved structure where every department head acted independently through their ministers. It was very chaotic. Games were being played by various parties. People [were] trying to sneak submissions past Cabinet through the back door. Foreign agencies providing aid funds would apply their own process. There were letters of agreement which would emerge later with the PM's signature. There were all sorts of interesting things happening, but it was clear to me that the public administration system was not assisting the government. (Igara, cited in Watt, 2006, p. 5)

In 1999, Igara's recommended structural reforms were incorporated into new legislation, *The Prime Minister and National Executive Council's Act*. These reforms led to the formation of the Central Agencies Coordinating Committee (CACC). "The CACC consisted of the heads of Treasury, Finance, Personnel Management, National Planning and Rural Development, and was chaired by the Chief Secretary" (Watt, 2006, p. 6). Igara's perspective on PNG's public management was that:

Most of the programs and the issues facing these programs were cross-sectoral. We would have economic issues entangled with how to best to manage relationships with donors. Fiscal policy and the deterioration of health services, infrastructure and public service integrity were linked. Then you would have natural disasters like the tsunamis in 1997 and 1998 and major issues like the crisis in Bougainville, which required several agencies' attention and consistent political leadership. (p. 6)

PNG has a mandated National Planning Framework operationalised in a cascading manner (see Figure 12). In 2013, the former Minister for National Planning and Monitoring, Charles Abel (DNPM, 2014) outlined that he considered the previous strategies and plans as “insufficiently strategic” and that the “strategic planning doctrine requires the identification of strategic issues within as well as external to the country, and developing appropriate policy responses to address them” (p. 18). The framing of the National Strategy for Responsible Sustainable Development (StaRS) in 2013 was this response DNPM, 2014, p.19).

In 2016, the planning framework was updated with the passing of the *PNG Planning and Monitoring Responsibility Act* (The Act). GoPNG outlines “the Act was to clarify the linkages and protocols in the planning process and establish a development framework” (DNPM, 2016, p. 3) with intent to complement other existing laws.

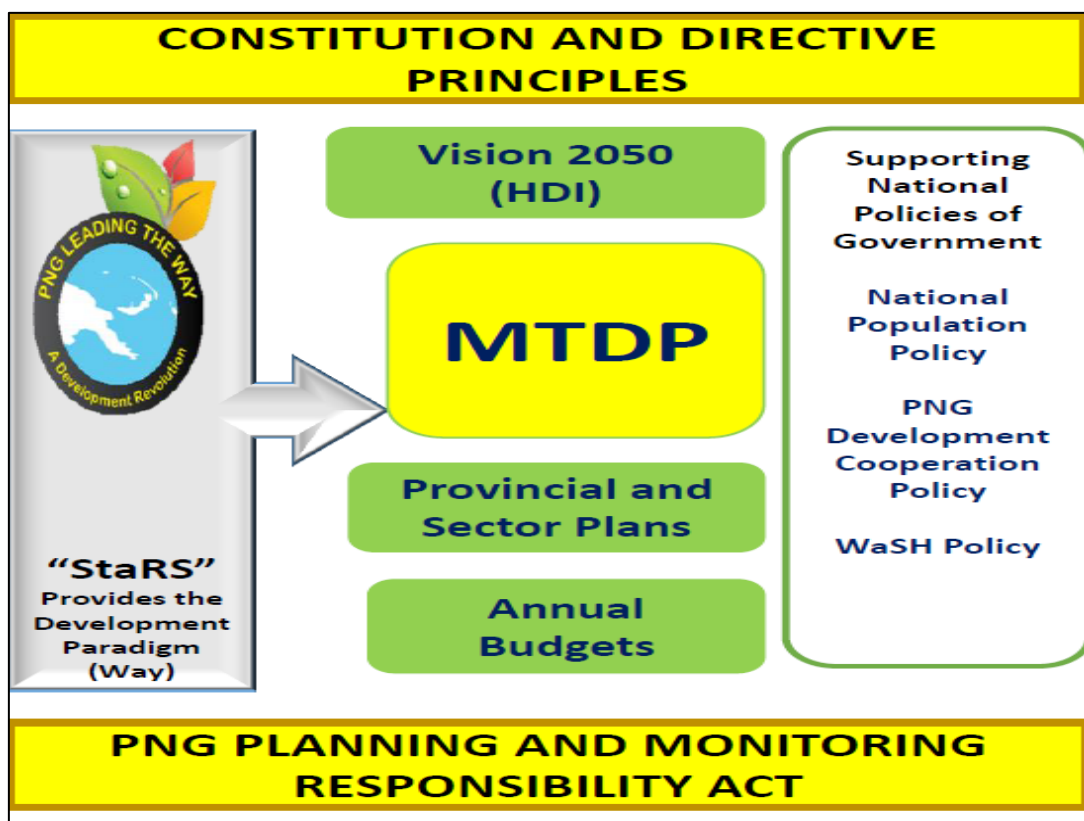


Figure 12. PNG National Planning Framework 2016 (DNPM, 2016, p.3).

The National Planning Framework (DNPM, 2016) consists of the following components (a-g):

- a. The constitution, which provides underlying directive principles.
- b. The NSRSD [National Strategy for Responsible Social Development] (StaRS) which prescribes the consequent development paradigm
- c. The Vision 2050 Human Development Index objective, and its subsequent reviews
- d. The MTDPs [Mid-term Development Plans]
- e. The Sector Plans
- f. The Annual Budgets relating to the implementation of the MTDP
- g. Any other documents made for such purpose. (p. 3)

The Development of the MTDPs coincides with the term of each Parliament (DNPM, 2014) which is five years. The application of the 2016 Act was outlined by DNPM (2016) where:

This Act does not legislate specific plans. It is intended to provide clarity and definition to a process for the benefit of all stakeholders (Government and Non-Government). It captures high-level principles and lower-level sectoral plans, international obligations like the new Sustainable Development Goals, together with the priorities of the Government of the day into the 5-year MTDP with clear indicators linked to annual budgets. (p.4)

Since Peter O Neil came into power in 2012, PNG governments have produced priority-setting documents such as the *Alotau Accords* (2012, 2016) which provided emphasis on the current government priorities alongside the current MTDP.

#### **4.3.3. Sustainable development and green growth**

As global attention shifted to more sustainable models of development (UN, 2015) incorporating and embedding sustainable development themes into PNG country development plans led to more sustainable country development goals. Embedding sustainable development principles into the MTDP 2016 became a focus for the government of PNG with development of StaRS (2014). Sustainable development was conceptually viewed in PNG comprising three intersecting dimensions of environment, society, and the economy. In addition, PNG shifted towards considering “a green growth economy” (DNPM-PNG, p. 18) based on the United Nations Environment Programme. A green economy was defined as “one that results in improved human well-being and social equity while significantly reducing environmental risks and ecological scarcities” (p. 18).

#### **4.3.4. Evaluation**

In 2016, the MTDP M&E framework was mandated at the central agency level within DNPM. In earlier versions of the MTDP, M&E frameworks and approaches were referred to but had no official mandate.

The MTDP M&E Framework (DNPM, 2016) included:

- i. A national reporting system that collects information on Government development activities to measure performance against the targets in the MTDP
- ii. Ensuring that results monitoring frameworks are produced & published annually including:
  - a. MTDP Pocketbook
  - b. Critical Activity Matrix
  - c. Capital Investment Report
  - d. Other relevant reports on the outcomes of the MTDP, and the annual budget.
- iii. Requirement that all levels of the Government & State bodies contribute to the implementation of the MTDP M&E Framework. (p. 5)

#### **4.4 Case research context**

This research focused on the theoretical perspectives and changes underpinning the architecture and use of strategic results frameworks, and management and government systems in two different development contexts. PNG was selected purposefully (Miles & Huberman, 1994) as a developing country case study as it utilised a national-led approach to country development, and PNG had agreed to incorporate OECD-DAC (2005) aid effectiveness principles in their national development planning approach. A fundamental aspect of the aid and development effectiveness principles advocated by the OECD-DAC (2005) was that developing countries would develop their own national strategic results frameworks and that donors would align to the country led development. This alignment was examined in my research.

PNG's use of successive development plans and associated results frameworks are centrally led and contained implementation intentions for line agencies under sectors. Increased use of MTDP's by GoPNG expects donors to align to PNG approaches for development, management and governance, and monitoring and evaluation. My research was undertaken in PNG between 2011 and 2015. There was a change of government in 2012, and this resulted in adjustments being made by GoPNG to focus on sustainable development based on the StaRS (2014) document.

There were significant differences between A/NZ and PNG in their approaches to planning country development. The principles that both countries base their approach were diametrically opposed. The research findings showed that in PNG a coordinated, centralised country approach is undertaken for development planning. DNPM used the MTDP (which aligned to the overarching *Vision 2050*, Department of Treasury, 2009) to link national development strategies and sector programmes such as the law and justice sector which received both national and development partner funding. PNG aligned sector and programmes to the centralised approach to planning and has some devolution of budget to provinces to meet their specific provincial priorities which cascaded from local-level government and community consultation.

With the *National Planning Framework* mandated in PNG, there were processes for inclusion and consultation of the provinces. Regional, youth and gender variations appeared included across sector planning. However, challenges were encountered in service delivery, measurement and accountability.

In comparison, A/NZ has no country development strategy or coordinated country plan. Under the *State Sector Act* (1998), central agencies devolved managerial responsibility to line agencies and their chief executives. These devolved reforms were in response to previously centralised approaches regarded as either too centralised or having an impact on the economy and communities from sudden significant policy changes such as the removal of agricultural subsidies. Moving to a devolved management approach in the late 1980s provided more flexible options. *Managing for Outcomes* initiatives and legislation changes included non-financial reporting requirements made to support the more devolved results-based approaches in A/NZ. It is these country contexts that provided the two contexts for my case study research in this thesis.

In Chapters Five, Six and Seven, the findings from my research were presented starting with those at the bottom of the updated Impact Model (Figure 7) – *ideas, interventions and frameworks*.



## Part B: Practice within the contextual layers

### Chapter 5: Ideas, interventions and frameworks layer

The following three chapters reported the analysis of the substantive and emergent research findings against the contextual layers of *ideas*, *interventions* and *frameworks* (*ideas*) (Chapter Five), *institution* (Chapter Six), and *wider infrastructure* (Chapter Seven). As shown in Figure 5, these contextual layers were adapted from Pawson's (2006) model and developed into a new typology as part of my theory-building research approach. This new typology was used in conjunction with the impact model research frame developed during the research scoping phase which was updated prior to the case study research (Figure 7). The findings that contributed towards development of my theory-building are highlighted. These three chapters addressed research question one: *'What evidence in A/NZ and PNG demonstrate 'what works for whom' in relation to 'results frameworks' and associated management systems?'*

The substantive findings for the *ideas* layer were drawn from three main sources: (1) the accountability document review (AR), (2) the rating of progress for impact model dimensions by stakeholders for A/NZ case study (RT), and (3) the qualitative findings (QF) from combined coding of the key informants in the scoping phase and the research participant interviews undertaken in PNG and A/NZ. A sample of the full QF was displayed in Appendix E, and a summary table of the QF themes in the three layers was provided in Appendix F. Once tabulated, the substantive findings for the *ideas* layer were further rated as: *showing progress* (and discussed in 5.3) or *requiring increased attention* (discussed in 5.4). Emergent concepts identified through these processes were described in section 5.5. A narrative of key findings was provided in each section to explain the data presented in the summary tables. A critical discussion of the substantive and emergent knowledge and practice concepts across all three findings chapters was provided in Chapter Eight.

#### 5.1 Aotearoa/New Zealand document review

My research examined the use of the strategic results frameworks and system changes in six key government agencies in the case study research. A sub-theme of the research was an examination of the inclusion for Māori and Pacific peoples within eight A/NZ public sector agencies (including two agencies focused on Māori and Pacific development). Documents used in the AR were chosen to represent the agency's public statements of accountability which include agency SOIs and annual reports. The findings for the A/NZ AR were displayed in Table 1, which were based on good practice concepts



identified during the literature review. These concepts were included in a good practice diagram that I developed (see Appendix B). The rating matrix data collection tool based on these good practice concepts was displayed in Appendix D, III.

**Table 1. A/NZ agency rating data.**

<b>Document analysis criteria (criteria included concepts identified from research literature review) (n=6)</b>	<b>Excellent</b>	<b>Good</b>	<b>Fair</b>	<b>Poor</b>	<b>Not evident</b>
1. Context analysis		1	4	1	
2. Needs assessment		1	4	1	
3. Higher-level framework in place	1	3	1	1	
4. Theory of change for higher framework is clear		4	1	1	
5. Baseline data is evident for results/outcomes			2	3	1
6. Needs of different population subsets (A/NZ Māori, Pasifika, etc.) are being worked towards		1		3	2
7. Trend analysis is based on robust data			2	3	1
8. Alignment of outputs from budgets to outcomes is clear		1		1	4
9. Aggregation of data from activities and programmes		1		1	4
10. Activities and programmes are aligned to higher framework			2	3	1
11. Intent of activities and programmes is clear		2	1	2	1
12. Theory of change/action for programmes is clear			1	4	1
13. Programme data is evident		2	3	1	
14. Use of and learning from evidence are clear		2	1	3	
15. Agency/organisation improvement is being tracked			3	3	

The review of the A/NZ accountability documents identified some ten issues of importance in my discussion.

First, the review showed that all six A/NZ agencies have a strategic framework in place (Criterion 3, Table 1). Their strategic frameworks highlighted intended outcomes and most displayed links to Better Public Service targets (used by the National-led Government from 2012 to early 2018). The evident use of strategic frameworks reflects a shift to an increased focus on outcomes under the *Managing for Outcomes* initiative (SSC, 2003), and mandated for inclusion in line agency accountability documents by the 2004 amendments to the Public Finance Act.

However, the extent to which levels of demand and need were determined through systematic evidence collection as the rationale for the funding allocation provided in the documents lacked detail, with most agencies ranking “Fair”. This raised questions about the use of data and level of evidence-based analyses for demand and need (Criterion 1 and 2, Table 1).

Only one out of six agencies achieved a “Good” level in relation to the evidence of services and interventions targeting the specific needs of Māori and Pacific peoples (Criterion 6). In five out of the six agencies, segmented population analyses were not outlined in the strategic intent, alignment of services and interventions, or in reporting documents. The principles of the Treaty of Waitangi included how the government needed to include the needs of Māori. Therefore, the implications of this finding were significant, given current levels of devolution to line agencies to enact such commitments. The evidence from the review suggested that to ensure systemic approaches across agencies with specific population groups, further attention was required.

The clarity of priorities and alignment of inputs to outcomes (Criterion 8) ranged from “Excellent” to “Poor” with four of six agencies displaying their intended outcomes. The theories of change for the frameworks (such as what the agency was working towards) (Criterion 12) were predominantly clear with four of six agencies rated at a “Good” level. This finding suggested that the agencies were cognisant of the need and inputs to align the intended outputs and outcomes.

More data for some of the business-as-usual services was evident at the output level than for interventions (Criterion 9). Both data sets were needed to display evidence-based progress towards the outcomes to demonstrate actual progress towards development results. The aggregation of data (Criterion 9) and connection of these data to the strategic frameworks was predominantly rated “Poor” with only one agency displaying aggregation of data from activities at a “Good” level. This meant that use of

aggregated data from activities was limited to support undertaking a contribution analysis from inputs and outputs to outcomes, and value for money assessments.

There was also a “Poor” to “Fair” use of baseline data (Criterion 5) and links to trend analyses (Criterion 7) which suggested use of evidential data to target services and interventions towards specific population groups was limited. This low use of data as evidence meant that actual progress towards the outcomes could not be substantiated.

The reporting at a trend level (Criterion 7) was clear with some agencies’ use of learning from activities (Criterion 14) also rating at “Good” to “Fair”. However, whether this learning translated to improved development results was unclear. Overall, in the A/NZ context, it appeared evidence for strategic alignment and contribution from business-as-usual activities and intervention programmes was limited, which suggested that feedback processes and adaptive management could be enhanced.

The demonstration of any contribution from activities to outcomes (Criterion 13) was overall either “Poor” or “Not Evident”, apart from one agency. The focus on input and output data in A/NZ agencies showed that an intended contribution from services and interventions to development goals lacked evidence and remains relatively unsubstantiated.

Three of the agencies rated “Fair” in terms of their tracking of improvements (Criterion 15), but three of them also rated as “Poor”. It was also not clear what improvements were required and what was happening to enhance agency effectiveness.

While no specific criteria were established to monitor gender sensitivity of policies and activities nor agencies’ attention to regional data, I noted that in A/NZ, gender-disaggregated data and regional data were not evident in four out of the six agencies. This finding suggested that gender differences and regional disparities were not being systematically considered, targeted or analysed across key A/NZ public service agencies.

Overall, this AR applied to the A/NZ accountability documents, showed that A/NZ agencies had made a shift to focus more on outcomes. However, a ‘line of sight’ from budget allocation and spending to results reporting remained predominantly at an output level. Overall, these findings suggested that public accountability documents were providing limited evidence of either alignment between need, demand and delivery of services or much evidence of increased tailoring of services and outcomes for specific population groups. These findings raised questions over the extent of outcome evidence in A/NZ and whether the needs of different population groups were considered in relation to what works for whom (RQ1).

## 5.2 Papua New Guinea document review

Documents from PNG were less easy to obtain, so the review was less systematic and primarily focused on the centralised guidance and planning accountability documents published per the PNG Planning Acts (most recently amended in 2016). These documents included the *Vision 2050*, *Development Strategic Plan 2030 (2009)*, *Strategy for Responsible Sustainable Development (StaRS) 2014*, and *Mid-term Development Plans (MTDP 1, 2010; MTDP 2, 2015)*. Most line agencies did not appear to have current (within one year) strategic documents and reports available. However, most line agencies had some level of plans (for three to five years) ranging from previous years to more recent. From the review of PNG accountability documents, some seven broad observational findings emerged.

First, the use of a mandated planning framework meant that the alignment of goals, plans (Criterion 3) and portfolio budget allocation was transparent at country and sector levels in PNG. Efforts were evident of increased awareness and understanding of sustainable development demonstrated by the development of *Strategy for Responsible Sustainable Development (StaRS, 2014)*. Sustainable development was an emergent theme in the PNG planning process and aligned with the emergence of the SDGs (UN, 2015).

Second, some baseline data was displayed by the DNPM against the *Mid-term Development Plan (MTDP 1)* (Criterion 5) and regional data by the Department of Provincial and Local Government Affairs (DPLGA). A Critical Activity Matrix containing a subset of MTDP indicators was used by DNPM for reporting at country and sector levels but was predominantly output focused with some limited outcome data (Criterion 9). Reporting regional data was not an impact model Criterion, but DPLGA reporting displayed variations in services between provinces which appeared linked to budget allocation.

The annual budget allocation was overseen centrally by DNPM divisions and Treasury. Each MTDP sector (health, education, justice and so on) had aligned goals, programme budgets, and processes for monitoring outputs. These planning documents provided transparency for development goal prioritisation and budget alignment to the PNG country framework.

The fourth finding was that civil society institutions such as churches and NGOs were used to deliver health and education services, and these inputs were recorded in the accountability documents. The PNG country framework and portfolio budget approach enabled the identification of donors and funding partners involved in the implementation of service delivery. The transport sector was one example which included the

Government of Japan working with the Government of PNG (GoPNG). The law and justice sector was another example where multiple funding partners (including PNG, A/NZ and Australian governments) aligned their contributions to the sector goals.

A centralised M&E approach was outlined in the two MTDP's (2010 and 2014) for coordinating data collection and reviewing progress. The central coordination of M&E activities and alignment of multiple development partners included GoPNG development projects, NGOs and other funding partners to the MTDP sectors (Criterion 10). However, apart from the Critical Activity Matrix, which was a data subset of the full plan, there was no detailed annual reporting evident against the two plans (Criterion 13). The monitoring of GoPNG development projects was undertaken by a monitoring division within DNPM, but progress data was not published.

The sixth finding, relating to operational and recurrent budget allocations was an area identified in the reporting documents that required further clarification. The performance of sectors at a provincial level with the centralised DNPM structure suggests there are execution and accountability challenges because of the limited availability and transparency of sector line agency reporting. While the MTDP planning framework appeared to enable input planning towards regions' needs and identification of intended outputs from interventions, there was little evidence in the documents that this takes place.

The inclusion of cross-cutting themes of gender, youth and sustainable development were displayed in the MTDP with evidence of regions contributing to the annual budget process. The mix of national and devolved authority for the regions demonstrated that the needs of sectors were being considered systematically within the portfolio-based country approach.

This document review showed the centralised PNG mandated approach did provide for country and sector level development goals and supported the inclusion of multiple funding partners, regions, youth and gender in planning processes. The use of a more centralised approach appeared to assist in a more systematic inclusion of different groups in the planning process. This finding concurs with the more collaborative 'new' national approaches outlined by Chimhowu et al. (2019). However, the limited accountability reporting and the need for expanded measurement of development results were two identified areas requiring further attention. The PNG accountability challenges were demonstrated by the limited auditing and published accountability documents displayed by PNG government departments. These limitations raised accountability questions between the central government and department allocations, spend and evidence of results.

### 5.3 Rating of impact model dimensions – ideas layer

The RT associated with the *ideas* layer by research participants during the A/NZ case study was provided in Table 2. This table mapped ten improvement dimensions against four simple standards that ranged from “Consistent progress evident” through to “unsure” (whether improvement is evident or not). The table synthesised a range of findings from the data sources described above and represented the researcher’s subjective view of what was reported across those findings. This rating was not undertaken for PNG due to the lighter level of case study research.

**Table 2. Rating table for impact model dimensions, A/NZ case study results.**

<b>Impact model dimensions for ideas, interventions, and frameworks layer</b> (taken from Figure 7). Research was undertaken 2012–2016. Assessment was based on 2014 (N=9).	Consistent progress evident	Some progress evident	No progress or change evident	Unsure
1. Increased use of needs analyses	1	4	3	1
2. Increased use of results frameworks for national development strategies and sector programmes	2	3	3	1
3. Increased focus on country/sector systems	3	3	1	2
4. Improved ownership of development	2	3	1	3
5. Improved clarity of sector development goals	2	4	3	0
6. Increased programme planning to meet priorities	4	3	1	1
7. Improved capacity in management	2	3	2	1
8. Improved capacity in M&E	1	4	4	0
9. Increased volume of aid flow aligned to national priorities	Non-applicable for A/NZ context			
10. Increased amount of aid is untied	0	1	0	8

Two observations derived from this Rating of Impact Model Dimensions by A/NZ research participants can be related back to the core criteria in Table 2.

Firstly, the response to Criterion 1 *Increased use of needs analysis* suggested there was some use of needs analysis in A/NZ. This finding pointed to some agencies

increasing their focus on more tailored services and interventions designed to enhance results for their customers. It may also indicate greater sophistication in the relationship between the demand and need for services and interventions, planning priorities and sector goals.

Second, Criterion 7, reflecting *Improved management capacity* also showed progress, yet it was unclear what aspects of management capacity were focused on for improvement. At the same time, it seemed that progress towards *Improved M&E capacity* (Criterion 8) was limited. This capacity lag has implications for managers and performance staff in measuring and reporting results using evidenced-based data. These findings highlight that more evaluative capability and capacity may be required by managers to enable more adaptive responses to changing contexts and needs of citizens, cognisant that there may be differences between ethnic groups and regions.

## 5.4 Substantive progress – ideas layer

Table 3 summarised the substantive conceptual findings indicating areas where substantive progress was evident in the *ideas* layer. This table (and the following findings tables 3, 4 & 5) were derived from a thematic analysis of the QF (Appendix F), then a mixed-method analysis was undertaken incorporating findings from the AR and RT. A narrative supported by the key QF for the *ideas* layer follows.

**Table 3. Substantive progress: ideas, interventions and frameworks layer.**

Concepts		Evidence
1	Increased use of results frameworks enhances strategic alignment, outcome prioritisation and communication.	Evidenced by RT, AR, and interviews with PNG and A/NZ managers, evaluation and performance management specialists, and development partners. This included the use of frameworks in planning and accountability documents in both PNG and A/NZ.
2	Use of theory-based evaluative approaches.	Evidenced in AR, RT, and qualitative data with more than three-quarters of research participants reported using 'programme theory' and 'theory of change' concepts (including architecture and use of results frameworks, theories of change, and contribution analysis) shifts the results focus towards outcomes to clarify intended development results.
3	Improved results measurement and funding alignment occurs from using results frameworks.	Shown by AR, RT, and qualitative data in MTDP and Critical Activity Matrix in PNG, and in A/NZ where results/outcomes frameworks became a requirement from 2016 in government agencies' SOIs and reporting documents.

#### **5.4.1. Results frameworks, strategic alignment and outcome prioritisation**

My findings noted differences in how results frameworks were displayed in PNG, A/NZ, and by funding partners. These differences appeared to relate in part to terminology and concepts not being systematically and internationally defined, which makes international comparability between planning documents challenging.

In PNG, the architecture of the MTDP framework had a tabular format for sector results frameworks (more in line with a log frame format originating from international development in the 1970s). Government of PNG (GoPNG) research participants said these approaches were both useful and a barrier to collaborating with development partners given the different international approaches to representation at government levels as well as differences in the ways multiple international funding partners chose to represent goals in framework documents. GoPNG increasingly using their MTDP as the overarching country framework with development partners and line agencies. This practice follows the guidance promoted by the OECD (2009e) of increasing the use of country systems and associated results frameworks (Binnendijk, 2000) to align multiple development partners towards a country's own development goals.

The mid-term development plan is aligned with all the sector plans and also with Treasury and the funding, and then linked to the recurrent budget and development budget, which is linked there. ...this builds on the midterm development strategy, which is 2006 to 2010 and ...and out of that came seven enabling priorities. (PNGM5)

In A/NZ, over half of the government agency research participants said they used results/outcome models of which some of the origins may be traced to Duignan's (2009) outcomes theory work. In addition, A/NZ Ministry of Foreign Affairs and Trade (MFAT) outlined requirements for incorporating diagrammatic results frameworks by development partners with interventions, and the Australian Department of Foreign Affairs and Trade (DFAT) incorporated theories of change as a key concept in the DFAT M&E Standards (2016).

Last year we released new policy and guidelines to support activity design which included our results frameworks. We're trying to finalise our policy and guidelines to support programme management, so that will include results frameworks at the programme level. We're also trying to go through a process... of just getting agreement to that highest level as well which is around the Statement of Intent. (A/NZM8)



However, from the mid-2000s, the architecture of frameworks in A/NZ was “more compliance-driven than out of strategic imperatives” (A/NZM3). Initially, multiple frameworks were developed within agencies for SOI documents alongside separate reporting, and M&E frameworks. Yet, this shifted to “being more around the key outcomes, simplification and not trying to explain everything, and measurability” (A/NZM13) as agency staff said they found it difficult to link inputs, outputs, outcomes and impacts, and demonstrate contribution due to complexity.

The movement [towards outcomes] is not only about having clear outcomes and having expenditure and various outputs and impact behind that, and it’s about achieving long-term social goals as well, social or economic goals. (A/NZE10)

My findings here highlight that the A/NZ government staff may be using the concept of complexity in this results-based environment as a way to potentially avoid aligning line agency activities and outlining their intended contribution to more strategic sector outcomes. The absence of systematically published sector and country plans by successive A/NZ governments’ highlights that the social and economic goals were not as transparent compared with PNG.

#### **5.4.2. Use of theory-based evaluative approaches**

Both managers and evaluators in A/NZ and PNG reported an expanded use of theory-based concepts with results frameworks. Theories of change (that is shown by multiple lined layers of results – inputs-outputs-outcomes-impacts) were a key concept which underpinned this increased use of theory-based evaluative approaches. These theories displayed the intended conversion from inputs, outputs to outcomes and impacts with different theories used in specific results areas. A sector example provided by one A/NZ manager was road transport where road safety and road transport infrastructure use different implementation theories of change inputs and outputs to work towards intended outcomes and impacts.

Outcomes theory of change links changes in state. You need to establish links between policy and outcomes and impacts and align goals, objectives and outcome targets. (A/NZM12)

Research participants from PNG and A/NZ suggested that programme theory using multiple theories of change as promoted by Rogers (2008b) were useful technical concepts to explain the multiple “strands influencing different outcomes – government, agency, provider, social service providers and whānau (family) – as different outcomes are sought at different levels” (A/NZE12). This approach may involve “unpacking what different levers are at different strands and with policy” (A/NZM5). A useful approach noted included using a “big diagram” mapping a “line of sight” (A/NZM14) from multiple

agency investments demonstrating alignment towards higher-level outcomes, and across agencies. This finding correlated with Mayne's (2001) contribution analysis approach where he advocated for pragmatic attribution by different agencies towards shared outputs and outcomes.

My findings showed that PNG used a centralised approach with results frameworks through DNPM whereas A/NZ government mandated agencies to identify agency-based goals and outcomes. In PNG, intended sector inputs with intended theories of change were more transparent as the MTDP was a published document covering five years (extending beyond the four-year electoral cycle) which participants considered provided "implementation continuity" (PNGM2). In comparison, A/NZ's intentions for achieving outcomes were outlined within agency documents aligned to a three-year electoral cycle. However, how the A/NZ government balances agency driven outcomes between sectors to support social, economic and environmental goals appeared unclear. This may be a factor contributing to inequalities (for different population groups, between regions and with environmental degradation such as decreased water quality caused by farming intensification).

The performance, our public sector, particular to management model itself is very agency centric and I think that is a weakness of the theoretical models that it doesn't explicitly have, a space for sector-based outcomes. (A/NZM20)

Research participants also said these identified gaps between strategic and service delivery/intervention levels can hinder how services and interventions contribute to strategic result areas as the links are not explicit. However, some research participants suggested increased use of sector-based approaches (for example, justice and social development) may be improving cross-agency collaboration. This emerging trend of using sector-based approaches in A/NZ is line with OECD (2009e) guidance.

The reason why we've started an initiative around social sector outcomes framework because what we were finding was depending on which business unit you talked to about the same outcomes in different ways. The idea was trying to look at what the indicators and outcomes were, trends and consistency, initially with moving more towards the social sector as well. (A/NZM12)

Research participants considered more outcome-focused leadership and detailed guidance from central agencies were two aspects which may enhance learning and performance across A/NZ agencies. A manager emphasised:

You do need clear goals and stability of goals over time with a good leadership team with your eye on those goals for a good five years to really built a solid

outcomes framework, embed it across the organisation and get everyone into understanding how their work contributes to it. (A/NZM6)

However, the role of evaluation was not detailed in A/NZ guidance which raised questions over the use of a systematic approach for measuring outcomes. Whereas in PNG, there was a centralised reference in guidance issued by DNPM, which outlined intentions of using evaluative approaches to measure development results.

#### **5.4.3. Improved results measurement and funding alignment**

My research findings suggested that the A/NZ state sector shifted towards an outcomes-focused approach with *Managing for Outcomes* mandated under the Public Finance Act 1989 (amended in 2004) but, research participants said, “moving from outputs to outcomes took time” (A/NZM16). It appeared that the passing Section 40 of the Financial Management Act (2004) may have contributed to an initial improvement in accountability reporting. In 2008, the revised Auditor General G4 accounting standard mandated reporting to include outputs, outcomes, impacts, and associated changes. Managers said reporting at these multiple levels of results was a fundamental change outlining what funds were spent on and contribution of the funded programmes to longer-term outcomes. Kauffmann et al. (2010) emphasised the importance of links between measuring results and accountability, and increased governance.

The SOI provides clarity from conversations and about what are government and ministerial priorities. Where the funding is going to go, and the frameworks assist dialogue and clarity. (A/NZM14)

However, another amendment was made in 2016 to provide greater flexibility in reporting for agencies. The measurement of actual results was an area identified for further research as this more recent amendment may have diluted accountability reporting for achieving outcomes. This finding raised concerns over whether A/NZ was sufficiently accountable to its citizens as highlighted in the Schick (1996) report. Gill (2008) outlined the reformed accountability relationships that were intended to improve accountability of A/NZ to the public. Yet, Dormer & Ward (2018) raised concerns over inclusion and accountability by A/NZ governments, which concurred with my findings.

Annual reports can [now] be just a list of initiatives but no mapping of direct contribution. [There] needs to be more joining up from intervention to outcomes frameworks for accountability, strategic alignment, reporting and accountability. [This involves] thinking more about outcomes correlation and agency accountability important. (A/NZM13)

The research literature review highlighted that contribution analysis was considered a key concept in evaluation for measuring development results and assessing

development effectiveness. However, in A/NZ research participants said alignment was not fully demonstrated from planning through to implementation, which made measuring results more difficult.

I think there's many people with passions in health and that's one of the reasons why this system measurement is really difficult because of that contribution thing and that makes it really hard to have that conversation about priorities and all that type of thing. (A/NZM9)

My research findings suggested that selecting relevant results for measurement was important to ensure associated dimensions of services and interventions met both "political and stakeholder goals" (PNGM7). Otherwise, tension may arise between "telling the contribution story and knowing what people have done" (A/NZM4). A solution put forward by an A/NZ manager suggested, "we need to structure our documents around frameworks, so it makes sense to have dialogue articulating the outcomes in the framework and show what you want to achieve" (A/NZM14). Research participants considered balancing between national, provider and whānau (family) levels to agree on priorities and collect relevant data.

what's the priority because it's not just an outcome here? We have multiple outcomes for many people distributed over the whole country and the population have different aspirations as well. So how does the Ministry's outcome thinking about outcomes encapsulate all of that in a balanced way? To some degree we have to reflect the values of the government of the day, but we also need to inform them in an ongoing way as well. (A/NZM2)

Durie (2004) emphasised that much more inclusion and consideration of Māori was required. Cram (2006) outlined that the impacts of colonisation are still impacting on Māori and that the government needs to consider ways to support Māori, which are more inclusive and collaborative. Research participants considered using frameworks as collaborative tools for engaging stakeholders may shift transactional management processes to being more transformational.

A use of results frameworks is positive as [you] can understand what and why [you are] doing. You need guidelines and support to get usage across the organisation. We need to ask within the agency – what do we need to do to [assess] whether it is achieving outcomes. (A/NZM6)

These findings may have implications for senior leadership teams and managers' capacity as using frameworks was regarded as specialist evaluative skills and "managers in A/NZ state sector tend to be generalists" (A/NZE2). However, Dahler-Larsen and

Boodhoo (2019) argue for more institutionalisation of evaluation and building evaluative capacity, which is supported by this research.

There is some existing evaluative capability inside agencies that can be utilised. People come from different backgrounds and [there are] also variations across agencies. [You] need more technical expertise to support managers and internally at the management level. (A/NZE2)

The evidence base suggested an increased focus on outcomes and non-financials by Audit A/NZ partially addressed identified reporting gaps. Moreover, over half of the managers and evaluators interviewed promoted intervention or investment logic models as technical tools to link programmes and projects to agency and sector frameworks. Research participants considered using models may improve agency and community participation and “link to longer-term outcomes and social and economic goals” (A/NZM14).

The research evidence highlighted that MFAT’s systematic evaluative approaches (introduced in 2012) used results models as tools aligned to strategic result areas, and this, when combined with training, appeared to enhance management processes and capability development. One manager reported using models had improved his confidence, technical skills, and assisted his collaboration with stakeholders through:

the awareness and understanding of the value of what we’re trying to do here – really great feedback on the results frameworks. People are, at the activity level, people are saying, ok, this is why we’re doing this and this. (A/NZM7)

#### **5.4.4. Summary**

The research analysis pointed to results frameworks being used in both PNG and A/NZ country approaches, albeit in different ways. Research participants said increased use of results frameworks (country and sector frameworks in PNG and agency results/outcomes framework and interventions models in A/NZ) assisted in clarifying and prioritising development goals and outcomes with stakeholders.

Overall, the findings suggested increased use of frameworks in both PNG and A/NZ may be improving transparency and alignment of inputs to outputs and outcomes. In A/NZ, the research analysis suggested an expanded audit function and support may be enabling a more evidenced-based management approach for development results. However, the analysis suggested that extended managerial and leadership capability may be required to adapt inputs and budgets within programmes, business units, and sectors in response to changing contexts and needs.

## 5.5 Substantive concepts requiring more attention

Table 4 summarised findings that show substantive concepts requiring more attention in the *ideas* layer. A narrative of these concepts for the *ideas* layer followed.

**Table 4. Substantive concept requiring more attention: ideas, interventions and frameworks layer**

Concepts		Evidence
1	Increased inclusion of population groups with targeted services and interventions.	Evidenced by AR, RT, and qualitative data. This includes assessing different population needs separately with tailored responses to meet cultural differences (such as with Whānau Ora health initiatives) and regional variations (as in PNG under Organic Law).
2	Using sector and systems approaches enables portfolio funding allocation and effectiveness assessments.	Evidenced by qualitative data from central and line agencies. This includes sector based approaches such as justice, social development and transport in A/NZ, and PNG's use of sectors in their MTDP.
3	Balancing centralisation and devolution of accountabilities is required.	Evidenced by the qualitative data from central and line agency managers and performance /evaluation specialists. PNG is mandated to consider regional variations under Organic Law and has included sustainable development as an underpinning theme in their national planning since 2010.
4	Increasing ownership and results alignment.	Evidence provided through RT, AR, and qualitative data. PNG's country plan displayed alignment of multiple development partners, whereas A/NZ displayed more siloed approaches between sectors, which requires more consideration to balance achievement of sector goals.
5	Integrating strategic planning, theory-based evaluative approaches and portfolio funding improves development effectiveness.	Evidenced by RT and qualitative data from central and line agencies. GoPNG displayed theories of change in each sector of their MTDPs and how they link to achieve country development goals (such as balancing tourism, agriculture and decreasing reliance on mineral extraction). This across-sector approach was not displayed by A/NZ in its more individualised sector and line agency planning.
6	Expanded leadership and managerial capability with integrated management functions and embedded evaluative processes may increase focus on achieving outcomes and impacts.	Shown through qualitative data from central agencies and line agency managers. A/NZ has placed importance of line agency leadership through separate initiatives facilitated by multiple central agencies. In PNG, the importance of provincial leadership and the value from capability initiatives were emphasised.

Concepts		Evidence
7	Expanded M&E capability and capacity for evaluation/performance specialists with training and development.	Evidenced by RT and qualitative data. Key areas in A/NZ included interpersonal attributes, cultural competency, and improved skills and understanding of key business and evaluation concepts. PNG passed a central agency mandate for evaluation in 2016 which involves capability development for central and line agencies, and provincial administration staff.
8	Using mixed methodology (quantitative and qualitative) for results measurement improves non-financial data collection, analysis, feedback, and accountability reporting	Shown through qualitative data from managers and performance/ evaluation specialists from central and line agencies. Both A/NZ and PNG research participants wanted greater use of qualitative data collection to provide more feedback on the effectiveness of government services and interventions.

### 5.5.1. Inclusion of population groups

My research analysis showed an “increased use of needs analyses” (Criterion 1) of different population subsets (particularly Māori and Pacific peoples) in A/NZ rated as either “Poor” or “Non-evident” in four out of eight agencies. Research participants suggested that state sector reporting for Māori was undertaken “only in part at a strategic level” (A/NZM22). The AR highlighted that only one out of six A/NZ agencies appeared to aggregate results systematically using segregated ethnicity and gender data from the business unit and intervention levels.

Research participants viewed approaches as “once over lightly but not involved in really driving change” (A/NZM14). The Treaty of Waitangi principles were emphasised in my findings, but participants said, “some do things, [but] communicating is not in a meaningful way and not enough time is taken into consideration and effect” (A/NZM8). Research participants emphasised that in five out of six key A/NZ agencies, the systematic inclusion of Māori and adequate consideration of their needs were not being undertaken, which breaches the Treaty of Waitangi principles. This view was supported by both Durie (2004) and Cram (2006).

The research analysis pointed to each A/NZ agency undertaking their operational approaches and processes under devolution of managerial responsibilities to chief executives. Over half of A/NZ research participants said an increased focus on Māori and Pacific development may be required with targeted strategies and interventions to assist a “shift from communication to influencing and measuring outcomes at the different levels” (A/NZE1). The findings highlighted that more systematic approaches and an enhanced role for A/NZ central agencies may be required to support agency managers more effectively, which may enable more equitable inputs and results. This finding

contrasted with GoPNG's more coordinated and centralised approach to national planning.

The findings suggested that setting targets for Māori and reporting progress at both higher and lower levels with “stretch” targets and specific interventions focusing efforts may enhance state sector effectiveness. It appeared that some target groups were “not being served” (A/NZM4). The evidence base highlighted that examining disparities for groups including Māori, Pacific and regions was required as research participants considered “averages hide” results (A/NZM6). Areas of good practice highlighted in A/NZ agencies where frameworks were used to target groups through “embedded” processes enhances practice included:

The sector-based approach with Whānau Ora (Family Health) really pulled sectors together. It used to be only at the strategic level and then a line of sight but was not followed up. (A/NZE10)

The research analysis suggested that specific agencies in A/NZ such the Māori Development Agency – Te Puni Kōkiri (TPK) and the Ministry of Pacific Affairs (MPA), and in PNG the Department of Provincial Local Government Affairs (DPLGA) for provinces, districts and wards may enhance “provision of advice and connect policy for service delivery and interventions” (A/NZM11). In PNG, research participants said DPLGA actively appeared to lead on approaches and tailoring inputs for regional priorities “supporting line agencies to work with the provinces directly to increase their capacity for service delivery and measuring results” (PNGM10). PNG research participants and documents outlined a systematic inclusive approach was mandated under the Organic Law (1997) which included “promoting equal opportunity and popular participation in government at all levels” (1997, p. i).

We basically do capacity building for provinces, districts and LLGs [Local level Government]. We also assist on monitoring, performance monitoring in terms of provinces performance for the preceding year as a requirement under the organic law. (PNGM10)

However, in both PNG and A/NZ, the analysis highlighted it was not clear how these focus agencies “link to networks and other agencies’ policies” (A/NZM11). In addition, a role for M&E across A/NZ agencies and within initiatives “is not clear” and there appeared to be “insufficient resources to measure results.” (A/NZM11). Other challenges noted in the A/NZ and PNG research analysis included coordination between country, sector, policies and regional plans, and that in A/NZ many administrative entities (such as city, district and regional councils, and district health boards) were not functionally aligned to national or regional development goals. Hughes and Smart (2012) raised similar



concerns in their views on the negative impacts from fragmented government agencies in A/NZ.

...you've got a distributed model. Greater clarity on roles is wanted to enhance sector performance. So, we can't think about sector performance without thinking about role clarity here – what is our role? We're mainly a policy advice agency. There are some programme things that we do but these people ...the councils – so they make decisions about what things get consented and what resources are allowed to be used. (A/NZM8)

My research analysis showed that in both PNG and A/NZ frameworks and plans were increasingly being used to support more inclusive collaboration as “we co-develop those [frameworks and plans] now” and ... it is around making the leaps from each step in the process from government outcome to frontline targets” (A/NZM6). However, the fragmentation (Ryan, 1996) evident in some sectors and services in A/NZ and the limited systematic inclusion and planning for different needs of key population groups and regional variations requires consideration, otherwise the inequalities evident between groups and regions are likely to remain unaddressed or may increase.

### **5.5.2. Using sector and systems approaches**

Research participants identified increased use of system approaches (including people in roles and components linking frameworks and business processes) as good practice. However, terminology appeared a barrier which may involve “getting clear what the terminology means for guiding people on approaches and priorities” (A/NZM10). Some agency staff said they find it “hard to understand where they fit in as the system is complex and involved” (A/NZM6). My analysis highlighted that using a system view with performance of an area people relate may assist thinking, as the “big picture and finance is important” (A/NZM4).

These findings indicated an increasing the use of systems approaches was wanted by participants. This was to look at “the parts that work and enable looking at the system overall” (A/NZM12). Research participants noted some performance variations were evident in components of services in multiple A/NZ sectors (such as justice, social development and transport). Integrating business units and having a ‘right mix’ targeting groups were advocated, including using Māori dimensions within the indicators with specific outcomes and targets. However, one manager reported:

it's not part of the ingrained culture of the organisation. You have whole lots of well-functioning individual units and have a poorly functioning system because they're not integrated or they're not the right mix to meet the population needs even though they are, as individual units, well-functioning. (A/NZM6)

These findings align with Williams' (2015) views on using system approaches for sectors which he considers enables a deeper understanding of key linkages and relationships. As part of my theory-building lens for this research, I observed that systems approaches appear important for more inclusive and transparent practices, and may contribute towards more equitable outcomes.

### **5.5.3. Balancing centralisation and devolution of accountabilities**

The research analysis highlighted that tensions exist in both PNG and A/NZ with balancing centralisation and devolution of accountabilities. It appeared that different approaches were used in each country context displaying both strengths and weaknesses in adapting to changing contexts and needs. The formalised planning approach in PNG appeared to include provincial and local level government and community inputs mandated under the Organic Law passed in 1997. These planning structures seemed to originate from colonial 'overseer' roles undertaken under Australian colonial rule. PNG research participants said after independence in 1975, PNG government struggled with service delivery and law and order, and they realised that new government structures and functions were needed from national through to local levels (PNGM2). These issues appeared to have led to the passing of Organic Law (1997) "to uphold constitutional values of sharing of wealth, community wellbeing rather than individual wellbeing and spiritual oneness with nature" (STaRS, 2014, p. 9). Yet, issues with agency service delivery remained.

Since the organic law in 1997, all agencies even the national agencies and provincial governments, provincial administrations don't have a corporate plan to guide them to deliver. Although they have development plans in the province, district plans, district development plans, but there was no corporate plan for them to deliver those development plans" (PNGM6).

This gap in planning appears to have led to additional support being provided by GoPNG with inputs from development partner funding for provinces and local-level governments:

we basically had to decentralise some of those functions to the lower-level governments, the provincial and local level governments. So that's service delivery would be effective and efficient. But ...unfortunately it didn't happen because the mechanisms or the systems that need to work were not properly established or built, the capacity were not built so that they could be able to deliver. And since 1995 and up until 2003 ...there was a need ...to reforecast trying to assist the provinces and districts and to defend their role as service

So GoPNG decided to provide additional support to provinces through:

...Provincial Performance Improvement Initiative, a programme that will assist really looking at public administration in those lower-level governments and build up capacities to the level. So in 2004–2006 when this initiative came about our usual plan was to assist provinces in coming up with their corporate plans so that they would be able to see where the gaps are and focus, on the areas that they think that they can be able to deliver and perform ... started with three pilot provinces, but ...have extended now to the rest of the country. (PNGM10)

In 2013, PNG incorporated sustainable development within its portfolio planning processes as Minister Abel considered that “previous strategies, plans and vision are not strategic enough to lead the country towards a responsible, sustainable and equitable future” (STaRS, p. 18). However, research participants said PNG has accountability issues for results and spending at local, provincial and national levels and that “there is an absence of an M&E framework [with the MTDP] which brings together the results from the different sectors” (PNGM5). Participants noted efforts by GoPNG where “currently that [M&E] is the key focus, and they’re thinking through how to do it and what is required” (PNGM5).

Development partners appeared to be using sector funding arrangements to account for their inputs alongside GoPNG systems in health, education and law and justice. One PNG based development partner outlined that this was not in line with World Bank and OECD-DAC guidance, which emphasised using country systems. As a response:

PNG is really looking at strengthening their financial management system. And that’s a major focus for them so then they can look at the funding coming into their financial management system, but they feel they understand [for] donors if they haven’t got their financial management strong in how it works – it’s really hard for donors to use. (PNGM6)

An example provided by one PNG based development partner was “the law and justice sector, a national coordinating mechanism where MFAT’s been working with Bougainville community policing is a major positive and also in the Highlands, so some progress [is being made] but using the parallel systems is disappointing” (PNGM5). The analysis showed that more capability and capacity were required for some PNG institutional systems.

Whereas in A/NZ, the findings highlighted that its devolved managerial mandate to chief executives to manage accordingly and identify results relied on leadership and managerial capability within agencies. A/NZ central agencies appeared to provide mainly transactional functions with the contested funding allocated through NZ Treasury. The

SSC employs chief executives and undertakes agency performance reviews. Research participants considered that:

Management for Outcomes lost focus and that measuring outcomes was [deemed] sufficient without telling contribution analysis story, which is important to demonstrate success. (A/NZM14)

However, from 2012, to address an identified performance gap, the findings showed that Audit A/NZ increased their leadership in measuring non-financial results. Research participants said they value Audit New Zealand's expanded leadership role, guidance and follow through promoting "intervention logics so [you] can say what business [you are] in and doing as need to audit outcomes" (A/NZM6). However, six out of eight A/NZ agencies reported using different evaluative and performance approaches within business units. In addition, one A/NZ evaluator emphasised that "don't necessarily think the government agencies have got a dedicated outcomes team...Their outcomes thinking is kind of predicated on what they've learnt through their experience in government, not necessarily [based] on some higher-level outcomes thinking" (A/NZE2).

My analysis showed over half of A/NZ research participants agreed that results measurement and analysis of progress within agencies does not occur systematically, or is done poorly and that some output and outcome results were "really varied and outcomes are really expressed in different ways. Quite often outcomes are fudged with outputs, and lower-level outcomes kind of resemble outputs" (A/NZE4). This QF was supported by the AR and RT findings where actual improvements using evidence to substantiate progress appeared limited.

In addition, the findings suggested that inclusion of providers, regional and community stakeholders were left to the discretion of each A/NZ agency. Yet the evidence indicated inclusion of representative population groups may be occurring in only one of out the six agencies (evidenced by RT and AR). My research analysis highlighted that balance between centralised planning processes, regional devolution, and accountability processes may require further attention in both A/NZ and PNG.

#### **5.5.4. Ownership and results alignment**

The Paris Declaration on Aid Effectiveness (OECD, 2005a) emphasised ownership of development as referring primarily to developing country governments' abilities to "exercise effective leadership over their development policies and strategies and co-ordinate development actions" (p. 3). In PNG, on the one hand, government agencies appeared to use the MTDP as a strategic portfolio framework for leading on country and sector planning and allocating funding and aligning partners' contributions (PNGE5). This more structured approach appeared to clarify sector goals and budget allocations

providing some transparency. On the other hand, in A/NZ there is a separate 'vote' budget allocation approach to agency funding. The voted amounts comprise a bulk funding mechanism for outputs from each agency. The funds are allocated to individual agencies by the NZ Treasury and this system tends to fragment (Ryan, 2011) overall development goals compared to the more sector-based approaches used in PNG.

In PNG, development partners were increasingly aligning inputs and funding to the PNG country plan (MTDP) and sector programmes. A PNG manager said "AusAid are going to try and get donors focused in four sectors – health, education, law and justice, transport because that's the areas of strength. And A/NZ in four, rather than focusing on a lot of sectors" (PNGM5). This approach appeared to support PNG's ownership of development where "donors base their overall support on partner countries' national development strategies, institutions and procedures" (World Bank, 2008, p. vii).

In addition, PNG research participants said the MTDP was increasingly being used as a country development planning and communication tool with agencies, NGOs and development partners to align and prioritise development results. This approach included "basically, looking at high impacts, customised tailored to meet the local context, gender, working together, how to meet the mid-term development goals 2015 reporting on both 2006 to 2010. Looking at customised targets with donors to work within" (PNGM6). Expectations of development partners to align with PNG's country framework and priorities appeared to be assisting to reduce implementation fragmentation in some sectors. A GoPNG manager noted:

...the big challenge is for PNG is the geographic and services in the provinces, how you work with provinces down to the districts, the service centres. The remoteness ... the feeling was that development partners only concentrate in some provinces and development effectiveness is for all people. We want to push more in to focus on hot spots but also use NGOs such as the churches, Lutheran churches in some of those harder to reach provinces/areas because they have capacity to do that. (PNGM5)

However, my findings showed that GoPNG wanted further alignment by development partners in measuring results as:

donors are doing their own evaluations but [we've] got to be able to get the information that comes in, so this is where the database about missions coming into the country, who's coming in, how that's going to benefit and how it's all linked together. So, PNG can prepare, and it can be used for that and systematic use. (PNGM5)

GoPNG outlined how they coordinate funders through a:

PNG coordinating, committee Paris Declaration, the plans for M&E, there's going to be an internal reflection meeting later in the year where basically you can jointly track progress against the targets, against the indicators, against the plans. The role of the database, the business plans, how they can be customised, so this is where the focus is for PNG this year. (PNGM10)

Moreover, since 2014, it appeared that DNPM reported annually against selected MTDP indicators using a Critical Activity Matrix as part of the PNG Planning and Monitoring Responsibility Act (2016) which were accountability requirements passed to increase transparency and accountability for ministers and key stakeholders. It appeared that a subset of sector MTDP results was used rather than the full MTDP and two PNG line agency managers questioned the operational use of some parts of the MTDP (PNGM8&9). In addition, the research analysis suggested that the DPLGA undertakes provincial monitoring predominantly at budget and output level in consultation with DNPM business units. However, PNG line agency managers considered that more limited collaboration occurs between DNPM and agencies than within sectors.

Whereas in A/NZ, a manager considered that objective setting "has been in large measure an emergent kind of a process, so ad-hoc requests and different work streams and requirements from government sector difficulties and emergencies" (A/NZM5). The absence of a country plan and limited sector development planning in A/NZ were highlighted in the research findings.

it would be untrue to say that there is a regular or that there has been a kind of a historical sense in which leadership of the organisation have talked together at length and decided that their job is to pursue the achievement of a set of objectives that has been widely consulted through the organisation and decided on". (A/NZM3)

The research findings highlighted the need to ensure "needs are more explicitly met" (A/NZE4) and "look at what you prioritise over others which is different from audit level when accounting for all things as need outcomes to become more stable, focused as getting what you are working to achieving and getting people on board is important" (A/NZM5).

My analysis noted that results-based approaches were viewed as useful, but skills and strengths varied as different units "do things in different ways and need to standardise approaches and how measure. The social sector tends to develop [their] own logics in [their] own unit, and people don't know about it" (A/NZM8). The role of managers, their capability and background in measuring outcomes were raised as issues

“as the tracking and measuring of programme is viewed as a manager’s responsibility” (A/NZM4).

Building skills of managers, programme staff, and leadership interests of training on the use of frameworks, and agreement on definitions appeared needed for “measuring and reporting at the right time to the right people” (A/NZM11). The frequency of reporting also required consideration as “getting info more frequently – weekly, monthly, daily so it [results] can be analysed more and get commentary around” (A/NZM6) with monthly reporting promoted. However:

the first step is pulling together framework so know what is important. There is value of different people being involved and [we] need to know what is important for monitoring and [what we] want to know more about so [you can] clarify what you do in monitoring and then [include] in an impact study. (A/NZM5)

In addition, my research pointed to the role of evaluation “changing where there used to be more external studies whereas the emphasis for evaluation studies is now more for organisational learning” (A/NZM20). The research findings highlighted that these shifts for M&E had implications for managers and leadership teams where their roles have expanded with the expectation of their use of technical evaluative tools. Gaps appeared evident in existing managerial M&E capability and capacity.

#### **5.5.5. Integrating strategic planning, theory-based evaluative approaches and portfolio-based funding**

In PNG, the findings showed that in the MTDP, sector goals and priorities were identified, and successive updates are undertaken. Data appeared to be used to identify selected priorities in iterative amendments to the MTDP with some development progress documented using updated baseline data within sectors. However, how updated priorities were set within sectors was unclear. Furthermore, no detailed annual reporting appeared undertaken against the full MTDP. PNG research participants considered challenges were compounded by geography, tribalism and a limited audit function of the PNG public sector (PNGM9). In comparison, the research showed that Audit A/NZ played a significant role in supporting line agency accountability (A/NZM1).

PNG’s country plans (MTDP 1, 2010; MTDP 2, 2012) indicated that PNG strives for balance between resource extraction and growing its economy through agriculture, tourism and other sectors aimed to meet needs of youth, employment, health, education, law and justice sector goals and service delivery (PNGE2). As part of a coordinated country and sector approach, GoPNG used a portfolio-based funding process to prioritise sector results and allocate budgets using the MTDP as a guide. PNG research participants outlined that the annual budget process involved a mandated process of

collating provincial government plans and budgets which cascaded up from local-level governments (LLGs) with community inputs. These priorities appeared to be then processed within sectors by line and central agencies, and the budget allocated annually. Both A/NZ and PNG managers considered that budget and prioritisation processes were more effective being “sector-driven rather than institutional driven” (A/NZM3). In A/NZ, the tension was evident between institutional-based funding and sector coordination.

Transport area tends to be, whilst they talk about a sector, it tends to be more institutional driven as opposed to sector driven. And quite often it's hard to get away from those outcome frameworks to be institutional driven because at the end of the day that's where the money's coming from; it's coming out through various agencies. (A/NZM14)

Two PNG research participants reported that provincial government administration had sector-focused personnel which were used as entry points into the provinces and for localised coordination and service delivery implementation. Under the Organic Law “provincial administrators as head of the province has to coordinate all services down at a provincial level” (PNGM10). Whereas the research highlighted that in A/NZ country development planning was predominantly undertaken at national level with two sectors tailoring services at a regional level (A/NZM10). An example provided by one manager emphasised that the Ministry of Social Development with its regional office locations seemed a positive way to implement regional priorities and service delivery. However, my findings suggested there was a need for increased sector-based tailored service delivery in the regions with targeted interventions cognisant of “ethnicity and age” (A/NZM11).

Moreover, A/NZ research participants highlighted how the justice sector was increasing its coordination and collaboration between agencies using a sector approach.

Those agencies are working quite hard to have an aligned sector strategy but then there are multiple ministers involved in there all with their own priorities so it's difficult to do. There's nothing in the model that prevents you from having a sector strategy and all working together on it but there's nothing in the model that forces you to do that either. (A/NZM21)

However, the way A/NZ funding allocation was made under a ‘vote’ allocation within each agency could inhibit inter-agency and regional collaboration as:

the short-term priorities and the agency-based accountability does tend to win out. You have to make a little separate space to get that sector base and get the agencies working together. But all the reporting in Parliament is agency-based, and the budget's all vote-based. (A/NZM8)



However, participants indicated that between 2012 and 2016, significant restructuring for functional realignment and role changes were undertaken within A/NZ line agencies (A/NZM6). Some agencies appeared to be changing their structures to work within themes (such as 'community' and 'education') which were "developed as a new constructive" to build relationships and provide governance.

We have four key theme priority areas: education, community and social enterprise, youth and skills, and languages and culture. Within each of those four theme areas, there are various activities happening between the relationships working with the communities and the policy team. (A/NZM11)

The findings showed though that across most A/NZ agencies a construct of operating in thematic groups was not yet widespread as most state sector agency personnel appeared to remain in functional groups such as policy, service delivery and research within agencies working with limited evaluation capacity and in communication "silos" (A/NZM4).

Whereas in PNG, with the MTDP sector approach and annual portfolio budget process, line agencies reported working in thematic areas rather than functional groups and budget was allocated through sectors with provincial inputs into the portfolio budget allocation processes. PNG research participants reported that thematic groups in line agencies supported more localised coordination as they were in direct contact with provincial government administrators and provincial service delivery teams.

However, funding accountability, geography and staff capacity constraints impacts on PNG provincial service delivery implementation. A PNG manager outlined:

We've started, I think we will not see something down the line if we can have good organisation whatever framework. It is people that make it work. It has to be people on the ground to – work in a way that brings things. (PNGM7)

In A/NZ, my findings showed that only three out of nine agencies demonstrated using baseline and trend data for reporting (AR, Criterion 5) which aligned to the higher-level results and funding allocation. Use of activity level data appeared "Limited" to Not Evident" (Criterion 9). Research participants emphasised a concept of value for money which "looks at outputs, costs, outcomes and impacts and makes a judgement call" (A/NZM12). However, there appeared to be "two paradigms operating in A/NZ, one to meet needs of government, and a second to meet needs of people on the ground" (A/NZM14). These findings pointed to a "mismatch" (A/NZE11) in A/NZ funding allocation processes where:

we have a scenario with the big outcomes that machinery of government wants to happen and programme and policies. There appears to be a mismatch – is it working? How do you know those pots of money are making a difference? We have people on ground saying, we're doing this, and we will have some money for that. It needs to meet in the middle. How are they actually contributing to those big outcomes on the ground – community organisations and providers?

Gill (2008) outlined the intent of the public sector reforms which was to improve accountability. Yet it appeared from this research that the devolved nature of A/NZ public sector planning and accountability may in effect be contributing to inequalities between different communities throughout A/NZ. Which agency has an overview of A/NZ country and sector progress (particularly social and environmental) appears unclear. The NZ Treasury coordinate the budget processes, yet accountability for outcomes – particularly, across the three sustainable development dimensions (UN, 2015). This finding directly contrasted with the central agency-led approach undertaken by PNG with their MTDP and their more recent mandate for evaluation to measure and report on progress. Given, the global shift towards 'new' national planning (Chimhowu et al., 2019), A/NZ remains in the 20% of the world's population that does not have a country plan. These findings and the concerns discussed below highlighted the importance of having a coordinated plan and the need for more transparent accountability from public spending for citizens.

In PNG and more recently in A/NZ (from late 2017 under a new government), funds were being paid directly to initiatives in regions which had raised accountability concerns in PNG, and in early 2019 concerns were also being raised in A/NZ. These concerns included “there's some, at the moment the funding's going straight to districts on, and the trouble is that, or with donors if you fund everything upfront there's no incentive” (PNGM10). A solution put forward by one PNG manager was:

to make sure things happen so basically planning [DNPM] want to get control back of the development budget and then you can look at progressively funding which donors would come under. And then that gives more control about controlling based on performance, on activities, outputs, outcomes and impacts (PNGM10).

#### **5.5.6. Expanding leadership and managerial capability**

The research analysis showed that an underpinning assumption for devolving managerial responsibility in A/NZ is that chief executives and agencies have sufficient contextual understanding, capability and capacity to plan and execute (SSC, 1998). However, research participants considered A/NZ government departments have “limited explicit direction and rely more on current term government and minister priorities”

(A/NZM6). Over the past decade, research participants considered agencies focused on improving management processes and meeting output class targets allocated under budgets rather than achieving “changes of state” (A/NZM6). In the A/NZ AR, only three out of nine agencies appeared to demonstrate a ‘line of sight’ from budget spend to outputs and outcomes (Criterion 8).

We’ve lost a lot of the intervention logic discussion due to decisions that we need to publish as short and concise document. Same with the Statement of Intent you will see we’ve had huge issues with Audit NZ around the way that we communicate specifically in our documents because we had to diverge from the official guidance because it would just end up being a too thick document. (A/NZM6)

Participants and agency documents highlighted that from 2014 A/NZ line agencies were shifting towards being more strategic and outcome-focused by improving accountability documents and non-financial information. Sector approaches and frameworks appeared to be increasingly used to align and overarch multiple A/NZ agencies (such as transport, justice, social development) and associated Crown entities. In addition, the analysis indicated that the direction and emphasis placed by Audit NZ appeared to contribute towards this shift. However, issues were highlighted:

we’ve never had, and we need an opportunity to actually develop the performance story to reflect a real-world scenario. And that will then make sure the linkage between what we do and the change we create will be much clearer, but I think it is not there yet. (A/NZM6)

My analysis pointed to development concepts underpinning planning and policy has changed over time for both PNG and A/NZ. From 2011, an initiative *Better Public Services* was embedded across the A/NZ state sector to enhance public service performance under the National-led Government. Indicators and targets were placed within specific agencies’ accountability and reporting requirements to support a more cohesive and productive focus across the state sector. This approach ceased with a change of government in late 2017. Instead, an increased focus on living standards (OECD, 2013) and wellbeing has emerged as underpinning conceptual development thinking in A/NZ central agencies, particularly from A/NZ Treasury.

In PNG, sustainable development concerns raised by the then Minister of National Planning and Monitoring resulted in the development of a sustainable development strategy (STaRS, 2014) to raise awareness of social, economic and environmental considerations in PNG’s development (PNGM2). Interestingly, Statistics A/NZ in 2008, produced sustainable development documentation, but an A/NZ research participant

made only one reference to this initial work. In both PNG and A/NZ state sector contexts, it appeared unclear how these underpinning concepts were operationalised and embedded into central and line agency planning and processes.

My analysis highlighted that both A/NZ and PNG managers and evaluators considered there was a need to expand their managerial capability. This was to enable a broader coverage of results-based management skills with increased use of measurement tools and extension to their understanding of key business and evaluation concepts.

Inside agencies, there is a lack of a business model on the business of outcomes not just on process but with a limited understanding of outcomes. Staff need to understand what business they are in, of converting taxpayer funds to impacts, and they need to understand this business model. (A/NZE4)

Leadership and managerial capability gaps were emphasised in the PNG and A/NZ research contexts which included the use of strategic frameworks and logic models as technical evaluative tools where:

having a logic model or model with shorter and longer-term outcomes means there is demand from policy to measure longer-term outcomes in the higher-level strategic space, which requires funding. If the programme theory is clear, then evidence sources can be identified. If intermediate outcomes are being achieved, it is more likely to achieve longer-term strategic outcomes by using models so then we can retest this later in a summative way later on. (A/NZM10)

However, the research analysis suggested that using models was predominantly regarded as the role of evaluators. In addition, A/NZ and PNG research participants considered there is insufficient evaluation capacity in government agencies.

It's around those country frameworks and the results assessments within those [frameworks]. It's a step in the right direction. Using partners, but that's about building the capability, so people are comfortable to do that and recognising where they need to kind of slightly also push and help everybody involved to move the next step—succinctly being able to articulate what we're actually achieving. It's a big challenge though, and it's not going to be solved overnight. (PNGDP)

Having evaluators (whether internally or externally located) develop stand-alone evaluation frameworks was viewed by one manager as fragmenting measurement approaches and constrains feedback on progress (A/NZM4). My analysis highlighted that leadership capability was important and that these capability gaps may be impacting

on resource allocation, tailoring of inputs for different groups and locations, and in assessing value for money. This highlighted that there may be management knowledge and practice gaps in using evaluative approaches for feedback loops to enable adaptive responses and organisational learning (Argyris, 1999).

Leadership needs to have a view and ensure incorporation of needs to generate work and outputs to contribute to subsidiary and up to the strategic level, linking in with strategic goals. The asset base needs to become more obvious regarding what is being done, what is contributing, and what are the results? (A/NZM10)

#### **5.5.7. Evaluative capacity for managers and evaluation specialists**

My analysis showed that an “Increased evaluation capacity” lagged an “Increased management capacity” (as evidenced in the RT). Research participants considered increased evaluation capacity and capability were required for both managers and evaluators in six areas.

First, there appeared a “very low base of technical skill for dealing with outcome and performance information” (A/NZM7). Evaluators appeared valued as they “can work from top-down to details which are needed to construct frameworks” (A/NZM6).

We need to develop skills and strengths in business units on how to evaluate and measure outcomes different business units do things in different ways – need to standardise approaches and how to measure. (A/NZM12)

Second, managers highlighted wanting to use more of a “system view to get clear about the object of change linking with social outcomes and the whole system” (A/NZM14) which included “trying to work towards common outcomes from community and provider to implementation at government level and looking at the links between what you’re delivering to what is being achieved and how you bring in innovation” (A/NZM3). One evaluator suggested that operations people can be “difficult to engage with to have a shared understanding” (A/NZE6), and therefore:

we need guidelines and support to get usage across organisation. We need to ask within agency what do we need to do to whether its achieving outcomes? A unified way of working sharing has to make the links of how what people does contributes to strategy and outcomes at higher level. (A/NZM5)

Third, the findings indicated that an understanding of the difference between research and evaluation is required “which is not the same” (A/NZE2) as “valuing the ability to squarely and rigorously assess how an organisation’s tracking overtime” was highlighted by senior internal evaluators (A/NZE2, supported by A/NZE3&7).

Fourth, increased clarification of the “intent of initiatives and what they are trying to achieve in different communities” was advocated as “iterative development evolves from community inputs” (A/NZE11).

Fifth, increased skills in context interpretation, quantitative and qualitative analysis may be required, and evaluative qualifications (A/NZE6&8).

Sixth, managers and evaluation/performance specialists emphasised personal attributes such as being “approachable so that people feel comfortable sharing and build up trust and be able to work together” (A/NZM7).

We need skills as a base inside agencies to share and have communities of practice. We don’t have people in-house who understand and who can evaluate effectiveness. (A/NZM12)

These findings concur with Ryan’s (2003, 2011) and Rothstein’s (cited in Dahler-Larsen et al., 2019) calls for more mature evaluation cultures within government agencies, which may assist in enabling more adaptive responses to enhance development effectiveness.

#### **5.5.8. Increased use of mixed methodology (quantitative and qualitative) and data collection systems**

My analysis pointed to enhancing non-financial data collection through increased use of mixed methodology (quantitative and qualitative) as “numbers and qualitative perceptions matter” (A/NZE2). In addition, in monitoring there was a need to include both as “context links the causes of the numbers” (A/NZM11).

You need to understand the context and assumptions so it’s replicable. You need to allow time to follow up to capture impacts down the line by taking a longer-term view. (PNGM4)

My research pointed to the high-level alignment of outcomes in A/NZ being used as “social outcomes frameworks” to link in A/NZ’s SOIs and as processes to align indicators for relevancy which can assist with measuring different perspectives” (A/NZE6). “Potentially for any outcome, there’s always the ways of measuring it and different perspectives on the outcomes as well. It’s kind of the whole multiple realities” (A/NZE6). However, the research highlighted these multiple result levels required different measurement approaches and qualitative questions which could improve value for money and development effectiveness assessments.

Now in the core business we look at effectiveness and efficiency, so that’s where you look for the value for money, is it achieving, is this agency able to demonstrate how it is achieving? How is it looking at its business and saying this

is the most efficient way to do it. So, we're looking for an agency demonstrating this. (A/NZM17)

My research showed that linking of data systems was wanted in both PNG and A/NZ as there were different levels of reporting within systems. Using systems approaches for integrating business processes and data was wanted as "institutions must have leadership and ownership within and have linked systems for M&E and financial" (PNGM4). There was an expanding interest to use data which included focusing first on a need for data and driving the demand, then moving into measurement, followed by knowing how data will be analysed and defining the success story considering different group variability. Participants emphasised, "measuring where influencing [with] baselines and changes over time as parts of data systems and cohesiveness for a sector with a line of sight" (A/NZE5).

The research evidence pointed to strengthening partnerships to collect data involving NGOs and using provincial data in both PNG and A/NZ. Access to data and collection in PNG from NGOs and churches was particularly important as they were providers of national education and health services in some communities.

Our partnerships with the churches, how the information from those programmes can be captured from because they're the programmes that effectively take us down to that provincial and district and local levels more so than some of the sort of core sectoral ones where we're kind of constrained by as far as the government system takes the monitoring within that sort of core agency. (PNGM10)

One A/NZ agency manager outlined how "clients do measurement for themselves" (A/NZM9). A change of mindset was also wanted in PNG as "some donors are working with agencies, understanding our culture, the organisation culture, how they can be able to assist to help us" (PNGM5).

So, they have changed completely, so they are now usually like working at our back, we lead, and they come and support us. They don't become leaders, they don't think for us, we think for ourselves, they help us, this is what we want, you help us in this area then we move forward. So, they've changed their mind-set. (PNGM5)

There appeared to be a growing focus on data collection processes and use of data by stakeholders in both PNG and A/NZ. Agencies were acknowledging efforts to improve their data collection and analysis capability and capacity while recognising that more coordination and support may be required. Data collection approaches and inclusion of regions, providers and communities were areas identified for further research.

These findings raised many questions for me as the researcher and evaluation practitioner over the intent and extent of the collection and use of segmented data for different population groups in A/NZ. I was aware (as an evaluation practitioner) that averaged data at a population level can hide disparities between different ethnic groups and regions.

### 5.5.9. Summary

The country plan and sector approaches used in PNG provided some goal and results' clarity, and the plan was used as a basis for dialogue by GoPNG and its development partners. This research highlighted that greater use of country and sector approaches in A/NZ may be required as the current agency-based approach appeared to fragment development goals and budgets. The findings highlighted that increasing participation in and alignment of sector M&E approaches by implementing organisations may enhance inclusion and an evidence base for substantiating progress towards PNG and A/NZ development results. Participants indicated this might require an increase in management and evaluation capacity and capability, and an application of more inclusive and systematic evaluative design and data collection approaches to provide evidence-based development results.

## 5.6 Emergent concepts requiring inclusion

The following section outlined the emergent findings that sit outside the main traditional focus of current evaluation theory associated with the *ideas, interventions and frameworks* layer. Table 5 provided a record of emergent conceptual findings for the *ideas* layer. A narrative of the findings was then provided under the conceptual areas.

**Table 5. Emergent concepts requiring inclusion: *ideas, interventions and frameworks* layer.**

Concepts		Evidence
1	Values of participation and collaboration are important for inclusion of ethnic, gender and regional groups, and requires explicit consideration.	Evidenced by qualitative data from over three-quarters of research participants in central and line agencies. PNG has mandated inclusion for regional needs. Whereas, A/NZ has limited approaches for systematic inclusion of ethnic group, gender and regional needs and variations, with no mandates apart from Treaty of Waitangi principles.
2	New paradigm for role and nature of strategic evaluation to enhance performance and learning.	Evidenced by qualitative findings in PNG and A/NZ. This included incorporating strategic planning, performance management, embedded theory-based evaluative design, alignment of policy, and performance frameworks, data and insights, reporting and accountability. PNG mandated for strategic level evaluation at central agency level. A/NZ has no mandate for evaluation at central or line agency levels.



Concepts		Evidence
3	Adaptive management and governance enable and supports practice.	Evidenced by qualitative data from managers and evaluation/performance specialists. PNG uses a portfolio-based planning and budget process which enabled managers within sectors to highlight their changing needs and requirements which are considered at country level, enabling annual adaptive responses. This is overseen by the national planning agency for PNG (DNPM). A/NZ has no central planning agency and NZ Treasury undertakes the annual budget processes. There appears to be a gap in national planning and accountability within A/NZ.

#### 5.6.1. Values of participation and collaboration enable inclusion

My research analysis showed there were mixed views on the nature and extent by which Māori and Pacific peoples were included in A/NZ government agency planning and governance (A/NZE1). One manager said “there is a Māori dimension too that we want to weave into some of this. So, some of those things have flavoured some of the indicators and things like that which is probably important to mention” (A/NZM6). Another manager outlined “it’s pretty much the iwi (tribe) themselves. So, it’s very much in a governance role. They’ve got a long way to go to set up their own structures and things how they work” (A/NZE1).

However, explicit consideration of what “participation” means in different cultures appeared required. This finding was particularly relevant for Māori and honouring the Treaty of Waitangi principles of “making sure Māori are part of society-social and economic and Māori succeeding as Māori” (A/NZE11). Research participants advocated “for a mechanistic linkage to make it happen” (A/NZM5).

Relationships were viewed as “central” to enabling practice (A/NZE4). However, as one participant noted it was the level of “participation” including plan together and undertake in a meaningful way” that results in “more impact and stronger relationships” (A/NZM5). Research participants considered this might “need a mandate to happen for community, policy, operational, performance and reporting for engaging at [the] outset as currently it is ad hoc” (A/NZM11).

We collaborate often with the Pacific teams that work in the bigger agencies as well. So, there’s the various levels of communication that go on. Primary push for many things but if we want to take relationships to another step and influence on something major and have that recognised by the other agencies is through our CE. (A/NZM11)

An A/NZ research participant reflected that participation relates to different cultures wanting the “same sort of things but doing it in slightly different ways” (A/NZM1). One

A/NZ manager put forward a solution of using a model approach for undertaking inclusion, participation and collaboration more systematically.

They are doing it a slightly different way, but at the end of the day, that's what they're doing. That's what's really driving them. And now when you go down and you have a look at the framework, that's basically what it looks like. They all [do it in] different ways. They all basically do that. Māori are interested in... that's why we talk about cultures in New Zealand. (A/NZM3)

Participant interviews concurred with the AR findings, which rated only one out of six key A/NZ agencies, including key population groups systematically in planning and implementation. Moreover, a central agency manager stated that “we don't look unless it's part of their core business” (A/NZM1). Another A/NZ manager outlined:

It depends on whether or not it's a priority and you're wanting to improve the outcome of those groups. So say if you are a District Health Board, they look at the outcomes of Pacific, Māori, European, Other – usually four or five graphs, and what they're seeing is Māori/Pacific Islanders outcomes are lagging behind others, and so they would set specific targets. (A/NZM8)

This research highlighted that an assumption of inclusion of key groups and participation was systematically happening appeared incorrect for honouring the Treaty of Waitangi principles under the devolved managerial mandate in A/NZ state sector. In addition, participation and inclusion were not undertaken in a meaningful way to influence and have impact.

So, we can say, and we do say and we have got evidence to say, is that the education system in A/NZ is actually really brilliant but there are some fundamental issues with how it runs that means that that impacts worse on Māori, Pasifika and kids with special education needs. (A/NZM9)

My analysis indicated that target setting for population groups also appeared insufficient as more inclusive approaches and processes require increased attention and/or explicit consideration.

So, we focus specifically on them, on the basis that the rest of the system's working pretty well. If we can make sure that those groups are doing really, really well then, the rest of the system will be doing brilliantly. So, it's kind of moving the system from good to great – in order to move the system from good to great we've actually got to target those groups that are just not actually being well served. (A/NZM9)

However, there appeared to be two other findings underpinning planning and implementation by A/NZ agencies that required explicit consideration. Firstly, a question arose of why some groups were disproportionally represented as 'not doing as well' within a 'brilliant' system and therefore required targeted interventions. This finding supported issues of who assesses effectiveness in a functioning system and its components. Also, a second finding pointed to questions on what is valued and who is included in designing the system and making the judgement calls on value' for whom where there may be issues of "power" and control evident within agencies.

There's a lot to do with who has power in a particular agency, who has the loudest voice, who has the ear of the CE, what's the CE's interest? I think those are really the questions that will influence those decisions and I think that's just the world as it is today. I would like to think that really good information that's put in front of people's noses in a timely fashion will help influence those decisions. (A/NZM11)

Other considerations raised by research participants included how decisions are made and for a role for evaluation in supporting a "democratic kind of more information" was advocated in making more "balanced decisions" cognisant of "power in that organisation" (A/NZM11).

Moreover, an issue of inclusion of communities arose pointing to a "mismatch between inclusion – community and providers with the government can learn from each other – need to bridge paradigms or change the paradigm so inclusive" (A/NZE11).

At this point in my research, participants when reflecting on the "accountability framework" in A/NZ suggested that a "value-laden idea of outcomes" may enhance processes and have more "impact" (A/NZM6).

The formal accountability framework in New Zealand, which covers both outputs and outcomes, has been implemented in terms of the management of a public organisation, particularly government departments, and we reached some conclusions that the management and use of outputs and processes had made a significant penetration. Probably the more value-laden ideas of outcomes had less impact on actual practice within management in government departments. (A/NZM6)

This finding suggested that "values are important and are often invisible" (A/NZM3) and that frameworks can help by making values more "explicit" so they "can be debated and contested" (A/NZM6). Moreover, using "top-down" approaches can also hinder understanding as:

where it becomes a bit weird is that then agencies are trying to use their top-down outcomes framework models to then measure whether what's happening on the ground through that little investment that's gone down to this agency or this community organisation that's doing a whole lot more than just this little bit of money that it's getting. (A/NZE11)

Whereas in PNG, the published country plan provided some transparency in development planning and intended sector goals. Yet, achieving goals and substantiating with evidenced-based results appeared "challenging" in both A/NZ and PNG contexts. Participants outlined the "subjective" nature of intended impacts such as wellbeing as you "can't force people to be socially well" (A/NZM12).

The findings showed that Audit NZ increased their support and guidance from 2012 and an A/NZ manager outlined:

The auditors want to see us as a government department talking about measurable achievements against the outcomes measures which we've set and to say whether we're doing well or poorly against those and to report quite rigorously against them and to apply an analytical blanket to what's going on, pull it together into a coherent story so that people can form an intelligent judgement about whether we're doing well or not. (A/NZM4)

Yet some A/NZ research participants raised an issue over who is involved in making judgements as:

value for money is just a judgement call that somebody makes after looking at your outputs and the cost of your outputs and the impacts, whether you're delivering your impacts, and where the outcome indicators, putting that together, are you moving forward, Yes, and does that feel like value for money. (A/NZM18)

PNG government and development partner participants also outlined issues of inclusion and participation in making judgements about development effectiveness. However, development partners appeared to display an awareness of the PNG development context.

On a country-level, how we actually go about implementing the results framework has required a new comprehensive aid policy framework. So, for the last nine months, our focus has been around results and outcomes and looking at what we can reasonably measure. What are the expectations from the agency side to judge performance compared to what is realistic and most useful in terms of the PNG context? Being able to make judgements about progressing towards the partnership development outcomes or not. What results do we actually need to

look at measuring to know whether we're on the right track to getting to where we want to be? (PNGDP)

This research highlighted that contextual understanding and collaboration were viewed as necessary at both “broad and specific levels to share knowledge and work together in complementary areas” (A/NZM5) for country development and with development partner relationships.

In posts [we] need to be operating at a higher level, particularly around these areas to be able to engage appropriately with our partners. That's a big part of our role is actually building that capability for the organisation so when they are meeting partners, they know what our policy is and areas we can work closer and harmonise with. (PNGDP)

My research highlighted that when allocating budgets agencies should “look at the needs of communities and dialogue with communities on what's happening, what's needed?” (A/NZE11). This could involve collaboration to:

work with other government departments to identify gaps where priorities are and to address gaps. [Its] important to know what other agencies are doing – to link on policy front on how better to enhance outcomes and work across sectors. And use evidence to show gaps with longitudinal studies. (A/NZE6)

Participants outlined this could include using service providers in “co-design” (A/NZE6) for services and interventions which “involves the strategic redirection of funds and in decisions on how funds are allocated by using outcome frameworks [from] ground up and aligning to a higher framework” (A/NZM10). An expanded role for programme managers was promoted by using more participatory approaches.

These findings raised issues for me of power and control in A/NZ by a predominant European population group and staff who may or may not be aware of different population group needs and the impacts of colonisation (Dwivedi et al., 2007). Innes and Booher (2018) emphasised the importance of including key population groups in collaborative planning which can lead to more inclusive and grounded discussions. However, the need for a decolonisation (Cram, 2016) process to ensure staff become more aware of their biases is being called for as part of becoming a postcolonial state (Gupta & Sharma, 2006). My analysis pointed to mixed approaches and processes being used for inclusion, participation and collaboration of key population groups. Increased attention and explicit consideration to these findings may be required to enhance development effectiveness and governance.

### **5.6.2. Shifting the paradigm for strategic evaluation**

The findings indicated that limited evaluative data appeared to be aggregated from programmes and activities within A/NZ agencies. This was evidenced by four of six agencies (in the AR) were rated “Non-Evident” for “Aggregation of data from activities and programmes” (Criterion 9). The QF highlighted this could be due to a need for evaluation as “little is done at present” (A/NZE5).

Very little evaluation is done, we have a small monitoring team and need to be thinking of setting evaluation design when setting policy objectives. We need internal and external debate and demand for evaluation as we need to know what is going on currently and information priorities. (A/NZM6)

A key finding highlighted in the analysis suggested that evaluation was regarded as “not at the strategic level as it’s perceived at being at the output level” (A/NZM1).

Evaluators and evaluation are at output level and does not have a reputation of being people who can connect strategy with evaluation. You need to understand what this means, not just for research but performance and the significance. Currently many evaluations if done at all, are not linked to bigger frameworks. This needs to be coordinated and a prioritisation. (A/NZM1)

In addition, research participants considered that A/NZ agencies were “not very mature about it [evaluation] – need to integrate as a way of working” (A/NZM6). Further comments on the nature and approach of evaluation highlighted “lack of integration of different functions” for programmes (A/NZM15).

Interview findings pointed to having a designated role for evaluation in A/NZ as there was not “a systematic methodological approach from strategy to implementers to providers as all different” (A/NZM14). Whereas “some systematic approach of implementers makes a big difference, but no systematic approach exists inside this agency and there is no oversight” (A/NZM12). One solution put forward by an A/NZ evaluator included “integrating research, evaluation and performance evolving to a more professional basis providing an evidence base and adapt according to size” (A/NZE3). My findings indicated this approach may enhance practice and results where “evaluation needs to be mode of operation...and need a common approach and [we] don’t have [that] at present and we’re dependent on some individuals” (A/NZM14).

Evaluation needs to be theory-based and link to the policy cycle – context-specific and value for money. Some process information is important, so you know what and why it is happening. (A/NZE4)

Linking “strategy, and planning team roles on frameworks and [for] accountability” (A/NZM3) was promoted as “including evaluators in working groups for frameworks makes it more robust” (A/NZE5). A placing of evaluation to “do evaluation for the efficacy of things, direction link with strategy and plan and have sensible measures and assess whether doing right things” (A/NZE2) could also enhance practice.

Further consideration for repositioning evaluation was seen by some research participants by “co-design linking management and evaluation” (A/NZM4) across the organisation, and for sharing information from a “learning perspective.” They suggested this may enable “evaluating strategic choices and adapting at strategic or lower levels” (A/NZM13).

Integrating evaluation into policy was also emphasised. One manager emphasised that “a solution to the problem was much more a sort of way about how we think and how we integrate evaluation into our decision making, outcomes of it and how we make it a constructive conversation” (A/NZM10). In addition, the cost of evaluation was given as another reason for it being “lightly done” due to relying more on external bodies to come in rather than [be] internally undertaken” (A/NZM5).

The research pointed to embedding evaluation within a system as it “is more than just basic monitoring. Two levels are required – activity level – providers tracking during implementation and strategic level using results and indicators” (A/NZE6). Also, extending capability is required and this can be acquired alongside “a system where we expect people to work in this way” (A/NZM21).

My research findings showed up areas for further research which included the links between policy development and sector planning. It was unclear from this research who sets the policy direction within sectors in A/NZ, and also at a country level. Increasingly, links between theories of change and policy are being promoted (Rolfe, 2019). In comparison, in PNG, DNPM staff develop country and sector level policy which interviewees outlined interfaces with line agency thematic policy development and implementation.

### **5.6.3. Adaptive management and governance**

My analysis highlighted that a focus on management processes may require more consideration and potentially a transformative shift to outcomes and impacts. This could be achieved by “not focusing on management process – outputs. Creating impacts and outcomes involves changing the state of things that is the business. Public servants need to know the business they are in converting outputs to outcomes and impacts” (A/NZM2). This may include:

clarifying the macro service model for government as current one is not working [and] requires transparency – outcome focused as need to have achieving outcome and evaluative dialogue and evidence in same space [as the] focus is on management rather than achieving outcomes – need to focus on the business you are in – creating impacts. (A/NZM14)

In addition, the findings highlighted being “iterative” and “identify things as you proceed as you will not know all indicators” (A/NZM4). Using “a top-down, bottom-up approach was emphasised including dialogue and workshop starting with a draft framework – key results outputs, outcomes, impacts and process” (A/NZM3). The research evidence showed that programme managers can “adjust based on evidence understanding the effectiveness of interventions” (A/NZE6). A glossary was needed “so all on same page and use common terms” (A/NZE4).

My research highlighted that adaptive processes and inclusive dialogue could use “living documents and make adjustments with two parts – collating information and dialogue with development partners” (PNGM4). And secondly, then “feeding it back to institutions for accountability learning and management processes” (PNGM4). A dialogue and workshop process, analysis and leadership were viewed as essential aspects with good communication and “contextual framing for key priorities” (PNGM6).

“Different perspectives and outcomes” were areas highlighted in the research and governance “on who can make the call for example – environment body, iwi groups as different views of units of outcomes of vested groups and owners and different perspectives and outcomes for stakeholders” (A/NZM6). The research highlighted that having a transparent evaluative culture including different perspectives in data sources was also “important as this is where the real outcomes are that you are trying to achieve” (A/NZE11). A PNG participant emphasised that “development effectiveness is for all people” (PNGM10). Therefore, “performance internally and perspectives of beneficiaries out to regional levels to compare region to region” (PNGM8) may be required.

These findings concurred with Hummelbrunner & Jones, (2013), Williams & Hummelbrunner, 2011) on the use of systems approaches and the more integrated and adaptive participatory approaches promoted by Shutt (2016). Further consideration of indigenous evaluation approaches such as Kaupapa Māori evaluation (Carlson, Moewaka Barnes, & McCreanor (2017) were highlighted as part of this theory building research as ways of including indigenous voices in country planning, management and evaluation.



## 5.7 Chapter summary

My findings developed an evidence base that pointed to several significant shifts that may need to occur if governments were to develop a more values-based management approach that, in turn, may enable a more integrated and adaptive governance response.

Overall, my analysis that underpins Chapter Five showed that increased use of results frameworks was evident in PNG and A/NZ, albeit in different ways. In PNG, the use of frameworks occurred at country and sector levels were, on the one hand, sustainable development in planning was emerging as an embedded concept. However, on the other hand, links from national to line agencies appeared limited. In A/NZ, however, agency results/outcomes frameworks remained focused at the agency level with some evidence that sectors were increasing collaboration and aligning outputs to outcomes. While there was emergent thinking on wellbeing and living standards in A/NZ, there was limited awareness of sustainable development compared with PNG's much more explicit focus on sustainable development.

An increased use of country and sector approaches in A/NZ may be required, as current agency-based approaches appeared to fragment focus on outcomes and goals. The results also highlighted that the country and sector approaches used in PNG provided clarity for GoPNG and development partners but mostly at the level of central government planning. The evidence also showed that increased participation in and alignment of strategic and sector-based M&E, including in implementing organisations, may enhance inclusion and sustainability for both PNG and A/NZ development results.

While both increased management and evaluation capacity and capability were advocated for in PNG and A/NZ, combined with more systematic evaluative approaches, this seemed to be at the level of advocacy only rather than for practice and implementation. The evidence, thus far, seemed to point towards the idea that incorporating explicit and agreed on values and integrating management functions and approaches may enable improved development effectiveness and governance.

Chapter Six follows, outlining findings for the *institution* layer. In Chapter Seven, findings for the *wider infrastructure* layer are provided. A critical discussion of the findings was then undertaken in Chapter Eight.

## Chapter 6: Institution layer

The second layer of analysis focused on the *institution* findings and addressed research question one: What evidence in A/NZ and PNG demonstrates what works for whom in relation to results frameworks and associated management systems?

Within this national planning, management and evaluation research, the *institution* layer applies to findings related to key line agencies (including health, education, justice, transport, social development, economic development, agriculture, and up to two other agencies (that is those with a cultural or regional focus). Findings from central agencies (such as Treasury or national planning in PNG (DNPM) were discussed in the wider infrastructure layer. The rationale for this is that central agencies' guidance and processes were designed at central agency level for a country and then associated activities are undertaken by line agency staff (managers and evaluators). However, this is not a definitive separation as in PNG, Department of Provincial Local Government Affairs (DPLGA) undertakes regional monitoring, and in A/NZ both Te Puni Kōkiri (A/NZ Māori Development Agency) (TPK) and Ministry of Pacific Peoples (MPP) respectively undertake monitoring of line agencies for their population groups.

### 6.1 Rating of impact model dimensions – institution layer

The RT associated with the *institution* layer by research participants during the A/NZ case study was provided in Table 6. This layer focused on structural issues relevant to frameworks and management systems. This table mapped ten dimensions against four simple standards that ranged from “Consistent progress evident” through to “Unsure” (whether improvement is evident or not). The table synthesised a range of findings from the data sources described above and represent the researcher's subjective view of what was reported across those findings. This rating was not undertaken for PNG due to the lighter level of case study research.

**Table 6. Rating table for impact model dimensions: institution layer, A/NZ case study results**

Impact model dimensions for <i>institution layer</i> (taken from Figure 9). Research undertaken 2012–2016 (N=9).	Consistent progress evident	Some progress evident	No progress or change evident	Unsure
1. Improved management processes	4	2	2	1
2. Improved M&E processes	1	4	4	0
3. Increased linking of priorities with expenditure framework and budgets	3	3	2	1
4. Improved operational development strategies	4	4	1	0
5. Improved results from activities and programmes	1	3	3	2
6. Enhanced evidence-based decision-making	1	5	3	0
7. Increased accountability	1	8	0	0

First, it seemed that *Improvements in M&E processes* (Criterion 2) showed only some progress evident. This area of improvement related back to the Accountability Document Review (AR) Criterion 4- *Theory of change/action for programmes is clear* and highlighted that processes connected to M&E need further work. This was further emphasised when compared with *Improved management processes* (Criterion 1) which appeared to show some consistent progress and supported that *Intent of activities and programmes is clear* from the AR, Criterion 11. This lag between improved management and M&E processes may have implications for the capability and use of strategic evaluation at country and sector levels and for evaluating progress towards development results. This finding aligns with Ryan's (2003, 2011) views of the need for more evaluation being undertaken in A/NZ, but growing this iteratively rather than following Australia's 'big bang' approach (Schick, 1998).

Secondly, *Improved use of operational strategies* (Criterion 4) appeared to have some and consistent progress evident but how these operational strategies linked to *Improved results from activities and programmes* (Criterion 5) appeared unclear. This may indicate an operational focus rather than strategic and transformational change intended as part of late 1990s public sector reform. However, *Improved results from activities and programmes* (Criterion 5) showed some progress or no progress evident. This could

have been due to issues in the measurement of development results and/or limited access to relevant and timely evidence-based data.

A third observation for *Enhanced evidenced-based decision-making* (Criterion 6) showed mixed progress where one third of participants (N=9) reported no progress and five out of nine participants reported some progress evident. This mixed progress may be due to capability of managers in using evidence and M&E processes not being in place to collect data. A mixed availability of relevant data was reflected in some or no progress evident rating of *Improved results from activities and programmes* (Criterion 5).

Fourth, *Increased accountability* (Criterion 7) was rated as showing some progress evident, yet it was not clear which areas showed increased accountability as it may have included more transactional management and operational processes as opposed to development results demonstrating transformational changes.

## 6.2 Substantive progress – institution layer

While the Rating Table 6 provided a snapshot of improvements in key dimensions identified from the literature review based on OECD-DAC guidance literature (2005) which was the entry point for this research, this next section focused on what were described as the substantive conceptual findings in both A/NZ and PNG from the participants. That were, as discussed in Chapter 5, findings that pointed to knowledge and practice ideas at a conceptual level, organised to assemble, on the one hand, those areas where progress was well underway and on the other, those that required further attention. The full table was provided in Appendix F, while the relevant excerpts from the table were highlighted here. One element (Table 7) provided the substantive conceptual findings showing

**Table 7. Substantive progress: institution layer**

Concepts		Evidence
1	Expanded auditing of non-financial information is improving focus within A/NZ agencies.	Evidenced by AR, RT, and qualitative data. This was mandated and supported by Audit New Zealand.

My research highlighted that an expanded role for Audit New Zealand, coupled with more systematic follow-through from them was improving general government agency focus on outcomes and performance, leading to an increase in transparency and accountability. This finding linked partially to the finding in Table 6 with *Improved management processes* (Criterion 1), but this may be affected by the lag in *Improved M&E processes* (Criterion 2). One A/NZ manager said: “we also have the annual audit, so they come in and they [are] now auditing and expressing an audit opinion on the

quality and the robustness of our systems and our performance measures and how they link towards the outcomes” (A/NZM4). A second A/NZ manager noted:

People haven’t articulated outcomes very well in the past and we’re getting better as a sector at articulating outcomes. I guess a lot of people have just put, almost, ‘impacts’ as opposed to ‘outcomes’ when they’re changing [their reporting] so [they] just make statements or targets for things as opposed to sort of, what I think of as an outcome. I generally think quite high-level ... almost a change in state. (A/NZ14)

However, there appeared mixed support within agencies for people to set targets as one A/NZ evaluator noted:

There is sometimes push back for people [not] wanting to stick their neck out on the line. If we want to set a target of crime will reduce by 5%, if we don’t achieve that it might be politically bad. .... [But] auditors are pushing hard to make them real. (A/NZE5)

The A/NZ audit process appeared “thorough” and occurring on an “annual basis” (A/NZ 6) and included changes to previous processes as “it helps to sort of sharpen the bench on what are we actually trying to achieve here, where are our priorities, where’s our direction? Everyone has to be on board” (A/NZM5). Overall, in the A/NZ context, this audit process was viewed as more aligned with auditing at A/NZ local government levels which had a longer history of auditing for non-financial information.

Whereas in PNG, however, only one agency-based research participant referred to PNG’s very limited undertaking of auditing at different government levels (PNGM9).

### **6.3 Substantive concepts requiring more attention**

Compared with the single progress area noted above, there were seven substantive conceptual findings requiring more attention for the *institution* layer, which were supported by evidence from the research (Table 8). These conceptual findings were identified from the analysis of the qualitative findings displayed in Appendix F. Each of these findings were highlighted in the following summary table, with a narrative of each area following.

**Table 8. Substantive conceptual areas requiring more attention: institution layer**

Concepts		Evidence
1	Increased focus and integration of strategy, inclusive evaluative design and adaptation for different contexts using systems and 'real world' business models as opposed to process models.	Evidenced in qualitative data, RT and AR. This includes linking and integrating planning, strategy, evaluation, policy, finance, communications, knowledge sharing and technology functions. This was evident in PNG where the MTDP contain sector goals, key outputs outcomes and targets. The operational role of evaluation in PNG was noted as underdevelopment.
2	Integrate strategic evaluation as a corporate function incorporating sector system evaluation, performance management and reporting to streamline for complexity.	Evidenced from qualitative data, RT and AR. This includes supporting activity level evaluative approaches that are integrated and embedded within thematic groups to enhance learning and accountability. Strategic level evaluation was mandated in PNG by DNPM in 2016. Both PNG and particularly A/NZ interviewees said that more detailed approaches were required such as the results-based approach implemented by MFAT in A/NZ. A/NZ has no mandate for evaluation with guidelines produced by one social agency (Superu, 2017) which has since been disestablished.
3	Increased integration between planning and policy as the balance of spend and value of policy over strategy was questioned.	Evidenced in qualitative data. An increased emphasis on strategic level was requested by research participants where policies align to strategies otherwise it was found to be adding to the complexity in A/NZ agencies.
4	Align resources and investment with outcomes and data systems to enable value for money and development effectiveness assessments.	Evidenced by AR, RT, and qualitative data. This includes more emphasis on results data, feedback, and learning analyses to ensure adequate rigour for non-financial management. This research found that the low level of relevant data that is available in PNG and A/NZ significantly limits learning, accountability, value for money and development effectiveness assessments, and governance.
5	Improved leadership, relationships, organisational culture and communication.	Evidenced by RT, AR and qualitative data. Increased strategic leadership by management teams, in both A/NZ and PNG line agencies is wanted to enhance relationships with sharing and learning as an organisational culture, using more governance structures across themes and programmes to support data sharing and oversee security.
6	Integrate management administration and process data, monitoring, evaluation, insights, reporting and learning as an increased focus on learning and accountability.	Evidence by QF. This included more emphasis in both A/NZ and PNG on results data, feedback, and learning analyses to ensure adequate rigour for non-financial management. This research found that a low level of relevant data is available, limiting learning, accountability, value for money and development effectiveness and governance.

Concepts		Evidence
7	Increased consideration of shared roles for central agency and provincial planning and by the public and private sectors and involving NGOs.	Evidenced by QF. This includes consideration of the planning and reporting cycle for agencies and roles for provinces in coordinating service delivery, monitoring and evaluating progress. PNG already uses a combination of government, private sector, development partner and NGOs to gather data and report on progress which is increasingly being aggregated to sector outcomes. A/NZ needs to use the data it already collects and develop approaches to aggregate this data up to strategic sector level outcomes. MFAT in A/NZ was noted as being successfully underway on aggregating shared data for key outcomes.

### **6.3.1. Increase focus on strategy and integrating performance management with inclusive adaptive evaluative approaches**

In relation to (1) the A/NZ research participants described an increasing focus on strategy and integrating performance management and inclusive, adaptive evaluative approaches for different contexts. Integrating planning, strategy, performance management and evaluation functions and linking these with policy, finance, communications, knowledge sharing, and technology functions was advocated by research participants as these functional areas currently overlap and create duplication. This emphasis for integration was evidenced in QF, RT and AR as the examples below demonstrate.

#### ***Strategy, evaluation and performance management***

A/NZ research participants emphasised a need to “drive strategy right through systems management and measurement ... it needs leadership” (A/NZM6) as “none of them [agencies] are very strong in strategy, monitoring and evaluation” (A/NZM17).

Dimensions of strategic planning is really important, not back-filling but what is needed and focus on needs analysis more. It is not an annual process and needs to be consultative as it relates to strategic thinking. It needs outcomes identified from discussion as a driving force, but there is tension with ministers’ short-term thinking. (A/NZM6)

In addition, “some departments do not have strategic plans” (A/NZM4). There was recognition of a “need [for an] overall picture and strategy and working knowledge ... [through] having a designated and agreed and well-articulated direction and end state” (A/NZM14).

The organisation needs to know where it's going and it needs to look out some time in the future beyond a political period, and it's really beyond that it's hopeless, that's just a short-term plan as far as I'm concerned. So those are the critical things that are needed. (A/NZM5)

My findings highlighted how “silos across strategy and reporting personnel teams creating a gap between strategy and reporting” where “the purpose of strategy is to go in a same direction” (A/NZM12). “Planning ahead” was emphasised as “auditing looks back, but we have to link plans, cost and funding as need to know how to get there and review regularly and refresh strategic goals” which requires “embedded planning for Māori not just a tag line” (A/NZE10). These findings concur with Mintzberg's (1994) view of strategic planning as he emphasises how the importance of strategy implementation is to actually achieve change for different stakeholders.

A/NZ research participants questioned the current role and scope for evaluation which was described as focusing on “discrete programmes” rather than “overall work of the agency” (A/NZM5). These participants argued there was a “need to look at big questions and not drill down as the role of evaluation is too narrow. You need to be able to navigate to service delivery levels and improve state feedback loops (A/NZM8). The importance of having a “systematic evaluation oversight” was advocated for learning and performing purposes (A/NZM9).

I don't think there's any systematic oversight of evaluation or monitoring of success within the Ministry. There are a vast number of different programmes underway, pilots that are being trialled, and a lot of them contracted through directly by the Ministry, other initiatives being undertaken by DHBs [district health boards] that we take an interest in. But every programme, every initiative, will have its own approach to evaluation or none and so far as I'm aware, there's no evaluation methodology for instance which is applied universally across the whole range of the Ministry's business (A/NZM9)

An issue of limited discretionary funding was highlighted as “most [funding] goes on staffing so need staffing to be responsive and people are limited in accountability” (A/NZE4). This qualitative finding was supported in the accountability document review. Research participants promoted “linking evaluation in business plans as there is informal linking but not structural functions and need to link performance reporting and evaluation. Evaluation is currently not part of the strategic corporate function as we have plans linked to the output plan but [these plans] need to be linked to outcomes” (A/NZE3).

Another consideration highlighted in the interviews was that “organisation culture affects evaluation as [it] comes from a risk perspective rather than learning” (A/NZE7).



There's genuinely still some kind of fear around letting go of control from our executive management perspective to allow evaluation to take place. Once people start realising what evaluation entails unless they come from a very scientific perspective which they normally don't – they come from a risk management and say let's just keep everything moving along as business as usual. They don't necessarily want evaluators to go down and identify areas even if you sell it as a performance improvement or a learning tool. (A/NZE4)

Linking planning and evaluation under programme management was promoted for A/NZ agencies to enhance results or “benefits” (a term used within A/NZ agency project management terminology). A solution put forward by one A/NZ manager included “have planning advisers in departments to link with the measurement of outcomes” (A/NZM8).

As one participant highlighted “building evaluation into business as a way of working, that considers all those wider contextual factors around sort of less compliance, less regulation, value for money so if it can be used as a tool in that context, it's potentially a really powerful lever” (A/NZE6). This included “what are we learning from some of those process elements as well as what's actually happening in terms of results as we potentially are losing some of that” (A/NZE6). Similarly, in PNG having evaluation alongside service delivery was also emphasised from “national down to provincial and local levels” (PNGM8). These findings support that a shift towards more mature evaluation cultures is required by managers. (Schick, 1996, Ryan 2011).

Moreover, under decentralisation in PNG, it appeared roles and boundaries required defining. In PNG, managers outlined limitations to progress in terms of geography and infrastructure as “PNG also has a big challenge because of the terrain, the geography it's very, very challenging .... Once we have the roads and bridges into rural remote communities then vehicles can travel there, government officers can travel easily. But with no good roads and bridges system in the country, it's a real challenge” (PNGM8) for data collection. They also suggested that administrative boundaries were a potential challenge. “We look at policies and legislations and things like that and we come to certain extent or certain boundary and then we say okay well you don't come in here because that's for your province or people in district” (PNGM8). A solution promoted by one participant emphasised the need for increased linkages and more collaboration which concurs with Shutt's (2016) more inclusive and participatory approaches.

We need the leaders and of course like me and others – we need to put our heads together with the provincial people in the district and work together to try and really change that level of service that is reaching our people in the communities. (PNGM8)

## ***Sector and system approaches and performance***

The participants pointed to using of “real world” sector and systems approaches to assist in “defining scope, identifying key components, linking with a line of sight”, roles and assessing contribution pragmatically (A/NZM9). This included:

Looking fairly hard at some of our frameworks and especially around how the system performs, how we understand system performance and measure it at that level and how we talk about it and communicate it and how that relates to sector performance which is obviously looking at the various components of the sector and how they’re performing and how they interact and below that and how that further relates right down to the provider level around how providers out in the sector contribute to the system performance. (A/NZM2)

Several research participants considered sector and systems approaches supported decentralisation and raised awareness of “influence on the front line” for service delivery (A/NZM12) which could increase accountability for results. Expanding devolution to the regions using a less prescriptive centralised and relational approach (Eyben, 2008) was advocated for by interviewees from this research.

The interesting thing is to see how the annual reports articulate with the statements of intents and because our level of influence on the front line in a sense what happens in classrooms and schools and kura [Māori language immersion school] is so devolved. So, the line-of-sight question is quite an interesting one in terms of how accountable, it’s not how accountable we are, but we need to be clear about the nature of our influence on the system such that we can be accountable for its results. (A/NZE6)

Responses from several participants suggested approaches for managing devolution and being accountable to citizens was displayed in the health sector with service delivery overseen by DHBs. This was a working example of regional service delivery in A/NZ as “the Ministry of Health works through a number of levers it can use to influence the sector and obviously it has the contracting of the DHBs and their role and what they need to deliver” (A/NZM14). In addition, these research participants considered:

[They] would find evaluation tools useful to use. I think [we are] getting there in terms of we’re not just accountable to Ministers. That’s how we’re working on our priority areas. We identify the activities, the inputs, the outputs, immediate short term, long-term desired outcomes. We’ve just finished doing that and we get together again as we have a coordination group. (A/NZM15)

Overall, the research highlighted that potentially by the integration of multiple functions using a sector systems approach that some of the duplication and gaps from discrete functional silos may be reduced which may enhance effectiveness.

### **6.3.2. Integrating strategic evaluation as a corporate function**

A second key conceptual finding at the institution level was the different ways evaluation was positioned within A/NZ agencies ranged from integrating evaluation into corporate functions to no organisational-wide approaches. Further consideration of the nature of strategic and programme evaluation functions, coordination, capability and alignment with business process were advocated. Participants considered the structures, functions, and approaches and how these impacted on participation, inclusion, transparency, accountability, and a democratic use of resources. A narrative followed with several extended quotes from participants' interviews illustrated this.

So, we're now developing and working with people on kind of broader evaluation strategies in those spaces so that's been a significant shift but at an overarching level we're not really contributing to the kind of the big intervention logic anyway. In other organisations, you see evaluation teams inside corporate functions aligned with personnel capability and working inside a team. (A/NZE3)

Research participants considered that greater knowledge of how resources were being used, results achieved for whom and systems improvements were required. A solution put forward by one A/NZ manager included:

I would like to see people with evaluation expertise working alongside people with performance management monitoring expertise, measurement expertise in a corporate function. But we don't have that and that may be an artefact of the way we're structured. I find it interesting that evaluation isn't explicitly identified as having a strategic corporate function. (A/NZM14)

Whereas in PNG, a M&E approach was part of a centralised function at DNPM with work underway led by central agencies levels to improve design and an evidence base for development results. However, greater support at the political level was wanted to increase transparency and accountability.

We need to have our political leaders must be leaders who have a vision for the people, leaders who are willing to work with the administration so that they can deliver services. We can't have politicians going you know with cheques and giving money to people, that's not what we want, we want politicians giving money to where it is supposed to – let the administration to deliver the service. I think that's one thing that we need to work on helping – even the elected members to understand – that we as a public bureaucracy have a role to play in

and they need to support us. They can't continue to say that we are not effective and efficient. I mean it's still in the past, we want to move on, we want to make changes, they need to support us. (PNGM5)

Efforts appeared underway to improve the functional alignment and nature of evaluation in A/NZ and PNG. This concurred with a lag of *Improved M&E processes* (Criterion 2) from *Improved management processes* (Criterion 1) evident in the rating table completed by research participants in A/NZ agencies. Further clarity of M&E processes and capability were advocated for by participants in both A/NZ and PNG. A solution put forward included:

integrate evaluation into a system as it's about evaluation being part of your mode of operation because otherwise, you don't have the ability to respond to it, to interact and how we integrate evaluation into our decision making, outcomes of it and how we make it a constructive conversation. (A/NZM10)

PNG and A/NZ research participants considered significant efforts were required to realign, restructure and integrate evaluation positioned at both corporate strategic level and programme level as M&E within business unit and internal coordination.

I would like to see the evaluation function inside this work and working more closely with management. I don't know that we've necessarily got the skill capability to that but in an ideal world think an evaluation function should be supporting people who support managers to do their jobs well in an organisational sense which is different from programme evaluation. (A/NZM21)

A multi-level combination of evaluative functions and activities may be required to provide integrated corporate functions, evidence-based results, share learning and enable adaptation to changing contexts and needs improving transparency for citizens and political stakeholders. However, an issue raised within A/NZ agencies included:

There are parts of the Ministry that need to have greater visibility with the Executive than [they] currently do. I think it's because they're responding constantly to ministerial queries and requests for new work and so on. It's easy to become focused on new policy development and new thinking about how to move the sector forward and the implementation of that in the work that the Ministry does in implementation and holding the sector accountable for delivering. (A/NZE4)

However, integrating evaluation and aligning with resourcing appeared to need "demand from executive management. If they don't ask the questions and if they're not

interested in the actual results, then it's very difficult for us and lower levels to actually advocate and create demand for evaluation" (A/NZE6).

This research raised questions whether the focus of agencies was on supporting management as opposed to focusing on their customers – the citizens.

[By] knowing the effectiveness of the programme [means] trying to drive towards [that] more clearly. That's an important function of the people who do this work in the organisation and in terms of supporting the leadership team to drive that hard. But again, I think that's also to do with the corporate function of the organisation which is to do with actually supporting management. (A/NZM20)

A role for central agencies was put forward as "what are the opportunities here around embedding your evaluative thinking in a system if you're driving the system off a managing for outcomes framework? In different jurisdictions that mediating layer is more or less controlled from the centre and in A/NZ of course we have much less line of sight in terms of frontline" (A/NZM12).

Moreover, relationships were viewed as important from national through to "beneficiaries of initiatives" and within agencies at different levels (A/NZM15).

Centrality of relationships in evaluation practice and relationships that we manage in our own agencies let alone with evaluators and the beneficiaries of initiatives and all those other relationships that we have to manage. But, actually, the internal ones are really important so it's just as important for me to develop really good relationships with senior management. (A/NZM15)

In addition, having an "internal culture to learn, hear challenges [along] with the good" was also regarded as part of an "internal evaluation function" (A/NZE6) which research participants considered needed to become more "open" to "soak it in and learn" (A/NZE4).

I don't think senior management were really open to hearing the bad news with the good news, they didn't want to hear anything bad. I think that's the same everywhere, but I think this organisation is going through a bit of a flux at the moment and I think there are ears that are open – for how can we do better? (A/NZE8)

In addition, the role of evidence information and role in democracy with using data in decision making was emphasised as "the whole idea of science and evidence and information and the rise of democracy all goes hand in hand since the enlightenment" (A/NZM8). This concept of democracy through information was linked to other questions raised earlier, including whose reality is it? What is valued? What does effective and

success mean? For whom? And other value-laden dimensions such as inclusion, and participation that is meaningful and have an impact. Research participants emphasised this required “mature thinking around it and certainly a mature practice about the way it integrates into or could integrate into its way of working” (A/NZM12).

We’re trying to think more holistically about information and its information which is really about knowledge to improve our understanding of how the whole thing’s hanging together in a number of dimensions. (A/NZM4)

These findings regarding the role of strategic evaluation concur with Dahler-Larsen (2012) who promoted evaluation as part of society and good governance. For me as an evaluation practitioner, the lack of a mandate for evaluation in A/NZ and the evident gaps in accountability and transparency over results-based government spending is problematic. I consider that the NZ and PNG public services require greater accountability and reporting of development results to its citizens so the extent of coverage, equity, and sustainability can be assessed. This view is supported by Dormer and Ward (2018) as a more democratic approach to improve accountability and public governance in A/NZ.

### ***Interface and integration of functions***

My analysis showed that the interface between strategy, evaluation and integration of functions such as performance management and reporting within institutions might require further attention in A/NZ as the “role of evaluation and strategy is still quite new and requires maturing” (A/NZM18). This could include integrating roles and functions within agencies.

There is a need to integrate processes how are those communities interacting – the financial community, the monitors, the evaluators, and finally the people who are engaging with and being the front face of the civic choice stuff? So, there’s a big supply, there’s a big connection between a whole bunch of professionals and really important roles there .., that are not quite integrating themselves. (A/NZM5)

Research participants outlined “there is a functional split between strategy and performance which causes challenges in alignment and understanding” (A/NZM14). “The reporting and strategy are undertaken in different areas of the organisation” (A/NZM12). Whereas, for PNG agencies, improvements were noted in planning and reporting processes.

We have become a little bit more focused in recent years – nearly three to four years ago we started. We have a corporate plan in place for the department and to accompany that corporate plan we have operational plans. The corporate plan

is more general in its orientation touching on policy and legislations – the bigger objectives. The operational plan specifically looks at outcomes that each of the sections or branches are supposed to be achieving in a year. (PNGM9)

One A/NZ agency evaluator outlined how their approach to address alignment involved “structural reform” (A/NZE8) by integrating strategic and activity level processes and results. This involved:

a review of the M&E function as part of the structural organisational change the [group] went through to ensure that it is better aligned with our strategic priority, delivering to your business needs etc., and basically building on the work to introducing results frameworks, monitoring systems so it’s a good foundation to build from. (A/NZE8)

However, the aim of the public sector reform (Dormer & Gill, 2010) was to increase the focus on achieving results and improve accountability and governance. However, my research raises questions over to what extent this was achieved. A key issue raised by research participants included a “need to define success, change of state – idea of actually how you define success” (A/NZM12).

If you’re doing outcomes you are trying to change society or you’re trying to change communities or you’re trying to change groups of people. Not just trying to describe a static situation, you’re trying to say well we’re trying to make something better, fitter, faster, increase, improve, optimise whatever it is. So, you actually have to say, there’s the end state and then you unpack what that is. (A/NZM2)

The relevancy of outcomes from national to communities for recipients of services was also questioned when:

you look at long-term council community plan and process and local government they are essentially defining outcomes at a community level. But then you’ve actually got central governments defining outcomes at a central level. Do these things relate, equate, are consistent, are coherent? So, if you have got outcomes being defined at a local level for a community and then outcomes for the nationality, at a national level being defined up here you’d expect there to be some integration. (A/NZM5)

In PNG, just the opposite occurred with localised approaches such as community learning and development which were used as:

mechanisms for community service delivery which actually mobilised community, also, we use those centres which we organise communities as well. These

centres have not become a whole of government kind of officially recognised service delivery mechanism as yet. The national research institute is using their information centres in the districts to disseminate information on their latest research into food crops and other foods and all that. Digital companies also are establishing several information centres throughout the districts. So, the focus of the department is all zeroing down to aligning programmes and projects – simply find them to a level that they can be what we call grounded in the communities, in the districts. (PNGM9)

Conversely, in A/NZ “more emphasis on department performance was being placed beyond chief executive and Crown entities and performance improvement framework” (A/NZM4). In particular:

you need to think about what people understand as an outcome. If somebody on the ground the beneficiary, the recipient of the service, what they think is an outcome and what a public service thinks is an outcome or the delivery person thinks is an outcome if they don’t have that same conceptual basis then you’re not really going to have people being able to co-produce outcomes for themselves. (A/NZE6)

Others noted though, for A/NZ, it appeared that outcomes were set at a national level, and co-production of outcomes was promoted to enhance impact. One challenge encountered in A/NZ was that “the output plan has limitations with linking outcomes” (A/NZM5) due to the accountability focus placed on outputs. In addition, difficulties in changing “output classes to align to outcomes” (A/NZM8) were barriers in linking inputs to outcomes where:

In our statement of intent for this year for example, we are linking a set of widgets up through our priorities to our outcomes we do link our outputs to our outcomes but in doing it we know that if we were really intent on drawing those links, we’d actually have to completely redo everything...our output classes and our output plan. (A/NZM9)

A solution put forward by one A/NZ manager outlined “if we want to have a real framework that is actually going to inform our performance starts at the bottom. I literally unpacked in a dimension logic model – all of the outputs the Ministry provides” (A/NZM13). This approach appeared to support earlier findings on using interventions/logic models at lower levels aligning to strategic results which MFAT used as an organisation-wide approach. In addition, “leadership teams need to be outcomes-focused” (A/NZM13).



how decisions are made using data and focusing on outcome is really important, not just accountability documents. I think the way the Ministry and the leadership team thinks is outcomes-focused, which is almost more important than whatever frameworks and things you've got. And if the statement of intent wasn't outcomes-focused, it wouldn't make any difference to most people in the organisation, but what we get measured on is the articulation of that in a set of predefined pretty poor accountability documents as opposed to actually is the way the organisation thinks and makes decisions based on outcomes. (A/NZE3)

### ***Decision-making***

My analysis showed eight out of nine A/NZ agency-based research participants (N=9) rate *Enhanced evidence-based decision-making* (Criterion 6) either as showing some progress (n=5) or no progress (n=3) in the RT. The key role of information in decision-making was highlighted.

It's about good decision making, that's really what it boils down to. The whole thing's about well-informed decisions ... it's interesting, when is it evaluation and when is it information for decision making? Evaluation to me is using information – deploying it back into the decision-making loop. (A/NZE5)

Research participants considered that integrating evaluation capacity could enhance evidenced-based decision-making by using “evaluation models and pools of experts or have people scattered throughout with community of practice. We need to link up internally, integrate evaluation into our decision-making outcomes and how we make it a constructive conversation” (A/NZM8).

Whereas in PNG, M&E was regarded as a centralised function, yet this function appeared fragmented between line agencies, DNPM monitoring, NGOs and donors undertaking various M&E activities. Research participants reported DNPM as the central planning agency was leading a design of a country-level M&E system (with development partner support) aligned with the MTDP (PNGM5, 6 & 7).

Given the devolved nature, A/NZ agencies varied approaches for roles, functions and linking of strategy and performance management evaluation, monitoring, and reporting. The AR showed there were limited systematic approaches in A/NZ or PNG agencies for evaluation (including monitoring) which may be due to an absence of centralised support and design. The interface of evaluation and other corporate functions required further attention and explicit consideration with “at least centralisation and coordination of the evaluation function in one specific point” (A/NZM10). Other jurisdictions such as Canada, South Africa and the USA have centralised guidance, country approaches and centralised oversight.

In framing of the evaluation strategy ... there needed to be an explicit focus on evaluative capability internally and for a period of time we did do some ad hoc organisation-wide professional development. But we never really kind of cracked the integration of that into the business planning cycle. (A/NZM18)

A stronger emphasis on “evidence of effectiveness” (A/NZM21) was promoted for both short term and long term. Expanded design, governance (with internal evaluation advisory group with some external expertise) and increased evaluative capability were areas identified requiring further attention alongside expanded “contextual understanding” (A/NZE2) including “driving hard for evidence-based culture for informed decision-making. Explicit consideration for regional and organisational integration with business processes and performance management” were emphasised (A/NZE6). A good practice example highlighted linking performance management and evaluation included:

Where performance measurement actually works was this one particular case in Australia where the senior management had lined up big meetings with evaluation staff and senior management and performance advisors or whatever. And they bought together not only the output reports which had all the indicator data in it, but also the results of any major evaluative work or research work that related to those same initiatives. To put it all on the table and have an actual real discussion about what does this mean for our performance? This can be an effective way for senior management to seriously consider how are we going and where are we at and what are we meant to do about it? (A/NZM18)

### **6.3.3. Increase integration between strategy, planning, policy, evaluation and finance**

In the third substantive area of the institution level indicating attention required, my analysis showed that linking between policy and performance management appeared disjointed. This was illustrated by an A/NZ manager as “policies and processes which don’t mean anything because they’re not used, they’re not driving performance. ...that is quite interesting in the people area, because you often find very exhausted performance management systems which people hate and they don’t use because they are just too unwieldy or they’re not trained or managed and its hard work some of this stuff particularly around poor performance. (A/NZM17)

One participant considered the “structure requires looking at as linkages are important” (A/NZM5) as different functions appeared located in separate groups such as policy, performance management and evaluation. An A/NZ evaluator outlined:

Planning across the policy team is limited at times and leads to multiple similar activities. Evaluation is important to link with policy as evaluation consists of several stages ranging from design and consultation through research through implementation (A/NZE9).

Integration between planning, policy and evaluation was promoted as separate functions “add to complexity” (A/NZM5). In addition, the value and balance of spend on policy over strategy was questioned “over the number of people in policy area – what is core business is that best use of funding?” What about if policy is all about what services shall be delivered? (A/NZM20). The research highlighted that integration may assist as “links with other teams are important to provide information such as finance and evaluation” (A/NZM4). My analysis pointed to an evident focus by A/NZ central agencies on the effectiveness of policy, mostly through undertaking a review within agencies and providing increased guidance with noted “improvements made in policy advice” (A/NZM21). However, this review noted that definitions and costs of policy still required substantiation as “policy is a trade off with increasing services and measuring outcomes” (A/NZM10). This review also led to questions on finance functions within agencies, intent to evaluate and use of data to assess progress and reforecast from learning.

In addition, consideration was wanted on “what’s panning out in [the] process, what does that mean, when do we evaluate and reforecast on what’s happened and what we’ve learned about whether our forecast was right or wrong?” (A/NZM2). Research participants suggested shifting focus to “what is happening”, and using “information in decision-making” (A/NZM21) was needed as:

My experience has been that that information is not used for decision making at all, or really it is just very patchy and I’m not sure that the blame for that lies entirely with poor evaluation it’s pretty common for new policy programmes to be set up without evaluation built in. (A/NZM21)

These observations pointed to the need for explicit consideration of integrating corporate functions (such as strategy, evaluation, policy, performance management and reporting) may be required to enhance development effectiveness.

Reflecting on these findings as an evaluation practitioner, I considered that the current paradigms and approaches of performance management and strategic evaluation appeared inadequate: there were a mismatch of terminology, gaps in the use of qualitative data in performance management (which can exclude insight into what is happening for different groups), and gaps in the explanations of how sector and country policy are formulated. This reflection caused a turning point in my research as I became aware that the current approaches to achieve more equitable and sustainable outcomes

may be insufficient. I considered that both performance management and the role of strategic evaluation may require reconfiguring and streamlining. I reflected that incorporating theory of change narratives (Rolfe, 2019) within strategic evaluation could enable more adaptive responses.

#### **6.3.4. Align resources and investment with outcomes and data systems**

Fourth, this research suggested that linking outcomes, planning, finance and evaluation was important to shift from being “finance-driven to strategy-driven” (A/NZM15). Several research participants noted mixed views on how agencies undertake their functions ranging from “haphazard” and “little planning” (A/NZM22) to “putting a lot of effort into aligning their outputs, their activities with that strategy” (A/NZE3).

Some agencies are reasonably investing a lot of effort into setting a clear strategic direction with the outcomes they want to achieve and putting a lot of effort into aligning their outputs, their activities with that strategy and then also thinking about their appropriate structure and the actual inputs that they’re purchasing and making sure those track all the way through. (A/NZE3)

However, an issue raised was that “we start hearing about the annual report when people come and knock on the door and say what can we say about how well we’ve done as it feels like a very disconnected process” (A/NZM4). Research participants advocated first “determining what you want to do”, then consider “how to fund it focusing on what outcomes and balance and link with annual reporting in the SOI” (A/NZM5) as “the high-level outcomes are so high level that the story, their outputs and then the inputs and the appropriations, it’s really unclear” (A/NZE3).

Quite a lot of organisations have their finance area, and they have sort of project and plans all controlled by the money available as opposed to determining what you want to do and then trying to work out how to fund it. (A/NZM22)

In PNG, a portfolio approach to funding sectors and agencies with the MTDP as a country and sector framework was used for budget allocation. GoPNG has expanded its use of MTDP, displaying increased national ownership and leadership with its country direction and alignment of donors expanding down to “province administration” (PNGM5).

The government to us is our major donor in this country. And since 2004 AusAID has changed its mind-set. It has now come in and working with agencies, understanding our culture, the organisation culture, how they can be able to assist in helping us. They have realised that they were not getting that impact. So, all they need to do is come and work with us, we know our own culture, how

we can be able to manage, we know better than them or donors, any donors for that matter. And I don't think donors have been as successful in the many cases because they think that they can come and impose on us, let me tell you, no you can't impose on us. You have to listen to us...you have to help us in what we think is right, where the lead we think is right – not what you think is right. I think that change has made a big difference. (PNGM5)

Another participant suggested in the development context of PNG; development partners were making progress on using country systems where:

they've even got their liaison officers in some of the provinces working with the provinces because when they need to see how the funding is going working in the province administration to see that whatever assistance is done correctly. So, we've seen a lot of change since 2004 and up until now we have seen a lot of changes in AusAid assistance for us. In fact, in our programme, AusAid is now using our systems and our processes, our accounting system for that matter. (PNGM6)

An increased focus on development results for people in PNG was also advocated:

to show PNG is doing something, women and children, so can demonstrate for every system kina, not just in one sector in PNG, what is being spent – what is happening. We need to commit funds to it. A reporting system should be at central agencies – National Planning (DNPM) [has] the development budget for the country – health, education – so making sure it's invested and accounted for. (PNGM4)

Whereas in A/NZ, there appeared to be a shift towards using a sector approach to provide clarity on the scope of the system, subcomponents and funding flow.

There's an education sector that there's kind of different layers, so you've got agencies, Crown entities and TEC and the Ministry of Education that collectively kind of, you've got Teachers' Council, NZQA, and then you've got sectors underneath it. So, you've got the schooling sector, which is two thousand, three hundred and whatever it is – semi-autonomous little agencies all by themselves. And then in early childhood you've got private and public, but the private sector completely unconnected... they're in it for profit. Most of the money goes to those right down at that level. (A/NZM4)

Yet, A/NZ research participants emphasised there were limited discretionary funds to influence results in A/NZ as most of the funding is allocated to paying public servants. This led to funding priorities being questioned within agencies.

the money that we have, or the agencies have to play with to change is quite small. So, the value for money question is really a very small question, a small amount of the 12 billion or whatever it is, can we actually make some value for money decisions about. (A/NZM5)

As part of the accountability focus, concepts of cost-effectiveness and value for money assessments were mandated by central agencies to assess agency approach and performance. In the Public Finance Act, entities were required to report by measures of cost-effectiveness. "When the Public Finance Act was amended back in 2004, the thinking was you needed bespoke cost-effectiveness measures" (A/NZM1). Noted challenges appeared to have led to a revised approach for cost-effectiveness.

Over time, we gave more thought to it along with Treasury and SSC. When you take a step back, if you've got a good performance framework and you've got good linkages and appropriate outcomes, impacts and outputs, then you don't actually need a bespoke measure as such of cost-effectiveness. You are able to demonstrate cost-effectiveness through your outcome frameworks. The assumption is there's an appropriate framework, and the value for money is kind of a related concept that brings in dimensions of cost as well. So, what are you getting exactly? (A/NZM8)

Moreover, from 2014 to 2015, A/NZ agencies made a shift towards value for money approach in A/NZ.

Value for money is built in our theoretical model because in theory you should be planning and setting your budget then looking at your results and your results that do it for the dollar spend it should be information, [and] in the theory of the system that information should come, be automatically generated. But I think in reality, it's not. (A/NZM20)

Internal politics inside agencies were highlighted by participants which may be impacting on transparency in A/NZ.

I think that it opens up too much transparency, that's five years' experience it's not a question that the capability is not out there. Again, it's a question of demand. There's a lot of clever people that can actually do this. As long as we don't have any major upsets, there won't be any further scrutiny. The problem obviously comes in when right like now we signal them for people, what's going to happen when you have to start rationalising your budgets? (A/NZE1)

In addition, an increased focus was highlighted on linking of financial and non-financial data "to be able to prove what you've spent your money on and what it did" (A/NZM13)

However, there were still issues in linking evaluation systematically with value for money assessments. One solution promoted by participants included staff aligning their efforts with “reprioritisation” using outcome data (A/NZM12):

Our front line they know they’re working towards an outcome even if they can’t tell you which one it is, for us actually it’s more and more important that we can actually reflect and acknowledge the work that they’re doing, and we can link it to an outcome because in the world of reprioritisation and showing you’re effective and efficient. (A/NZM12)

A systematic approach and using programme theory were two concepts promoted by research participants for aligning financial and non-financial information, aggregating development results and analysing in sectors and thematic areas. An increased focus undertaken by Audit A/NZ on outcomes was evident in this research. This finding reinforced how centralised guidance for A/NZ agencies could improve accountability and transparency, which was in line with Rolfe’s (2019) approaches of integrating theory of change narratives alongside results-based frameworks. Focusing on value for money through measuring outcomes was promoted by some A/NZ research participants.

Audit A/NZ and Office of Auditor General think it’s really important and they reinforce it’s just demonstrating value for money and demonstrating why you exist to the public otherwise why should we be here if we’re not adding any value and we’re not delivering any outcomes? And I think that’s a challenge to make them actually meaningful and adding value, so not just the easily achievable ones so we can all get our bonus this year. (A/NZM5)

There appeared to be some improvements in A/NZ for assessing value for money, but the lack of outcome data measurement remained an issue. Expanding the functions of central agencies was promoted by research participants to support more systematic approaches for outcome measurement

### ***Outcome measurement information gaps***

My analysis highlighted that to assess value for money and effectiveness it required data information at the “top, middle and bottom” and there was “still an information gap in the middle” (A/NZM8) where:

at the very highest level of outcome things, we actually know quite a bit of information. We know about children’s literacy rates, who doesn’t get well served by the justice system or the education system or the health system or whatever. So, at that very highest level we know quite a lot. And then at the financial level,

at the cost of services, we know quite a lot, and in between, we actually know very little. (A/NZM15)

This identified information gap included:

we know very little about what results from services drive, what impact they have. We know very little about the quality and efficient production of services. We know very little about the costs in their relationship to services. So, it's kind of like we've got good information at two extreme points and weak bridges in the ladder in between. (A/NZM5)

In addition, an issue of "too much transaction information" was highlighted by some:

There's too much built into the transitory, transactional information and not enough – what are we using it to try and understand them to do? So, there's a whole demand-side problem about – if we had all this information what might we do with it, and why is there a demand-side problem? (A/NZM15)

Context and progress on value for money and effectiveness appeared to have its origins from A/NZ's 1980s structural reforms of state sector and devolution for management to agencies.

I do think that an important principle that has entered into the system which sees the pushing of information up to Parliament but leaving room for the Executive to govern and manage – so be responsible for their choices – as opposed to imposing a whole lot of structure and rigidity downward. There were a sea of reforms that have similar things about them, they share an emphasis on performance measurement and performance information. Alongside that they've introduced financial accounting disciplines, so we've had the AG4 on financial information and performance information, and whether we get the value out of it is a different thing, but it's been easy to implement, manage, control and at a base level of quality of the information, it's quite stable and quite recognised and quite accepted. (A/NZM15)

Yet from this research, there are questions in both A/NZ and PNG contexts over the role of governance at country and sector levels, and the managerial role of agencies, and sustainability of services (A/NZM8), within and between government agencies. These issues together with financing of country plans were recognised by Chimhowu et al. (2019) and the OECD (2019).

Research participants also emphasised the importance of a feedback loop.

Because evaluation is not demanded by ministers and so the effort goes in to all the setting the policy up and no thought's then gone into what do you actually do



in two years' time when this programme has been running for a little while to see if it's any good. ...My observation is a lot of agencies have struggled to get back into that, what are we actually getting for the dollars we're spending mind- set? (A/NZM5)

To address these mixed views on information and performance gaps, participants considered a "function inside agencies to get rigorous information" may be required as "a way of building the rigor for our external accountability, we need rigor regardless of whether it's beneficial to us as an organisation or not but we need it" (A/NZE2). In addition, the allocation of spend on non-financial and performance information compared to financial information was raised as another area consideration.

One thing that amuses me is people will sometimes talk about the compliance cost of performance information, I say, Goodness, in your organisations I don't see vast tens of people dedicated to collecting and managing your performance information processes and transactions, but I see that for your financial information and no-one seems to notice. So, there's a high level of acceptance of the need to manage and record and measure... the units of money, and less so on the performance. I don't know why that is. (A/NZM20)

Overall, research participants emphasised that an increased linking of priorities with expenditure framework and budgets was required including aligning resources and investment with outcomes and data systems. This linking of financial and non-financial data may enable improved value for money and outcomes assessments and ultimately enhance development effectiveness.

#### **6.3.5. Improved leadership, relationships, organisational culture and communication**

In the fifth substantive area at the institution level indicating attention required, participants promoted increased strategic leadership by management teams, enhanced relationships with sharing and learning as organisational culture, using more governance structures across themes and programmes to support data sharing and to oversee security. This was evidenced by the accountability document review (AR), rating table (RT) by A/NZ participants, and the qualitative findings (QF) highlighted in the following section.

Expanding leadership was emphasised "from top-down and is inclusive so people know where they contribute and why" (A/NZM4). The "role of the Senior Executive team is important in discussing priorities" using an "outcomes" focus (A/NZM8). *Improved management processes* (Criterion 1) were noted in the RT, yet it appeared from the QF that the focus appeared on "project management" rather than "outcomes" (A/NZM14).

An A/NZ manager said “not focus on process – as there’s too much on process and a lot of emphasis on project management, rather than outcomes. We need to work at [the] strategic level but go down to details” (A/NZM14).

Other A/NZ participants pointed to an increased linking of goals, objectives and outcomes by senior leadership was required as the “chief executive (CE) is an outcomes manager” (A/NZE4) with more focus across “results areas rather than on the “management of people” (A/NZM10). The external review process by States Services Commission uses the Performance Improvement Framework (PIF) as an “interface is important between CE, PIF and organisational management linking strategic goals and monitoring and evaluation” (A/NZM20). Yet evaluation was specified within this interface of functions.

In addition, there was a “need to have [the] focus on outcomes not outputs and use business planning process as an adaptive tool rather than on project management but need to clarify accountabilities” (A/NZM20). Greater emphasis on accountability, reporting and governance appears required as staff “need managers to know and care that results matter” (A/NZM8). The role of “governance groups” was also emphasised as “people in management must be accountable for outcomes. Cost out results and opportunities for what need to know more about and have accountability governance groups driving performance” (A/NZM8).

Challenges raised by A/NZ research participants included the “system is unwieldy, and people are not trained and have people so within themselves and their area, and it is so hard to own it and drive and be accountable” (A/NZM8). Research participants promoted “using business plan processes to update and adapt off SOI and outputs plan. SOI comes down to output plans so you can make a decision to stop doing some things” (A/NZM6). Other considerations included having “corporate governance at the fore of reporting and using measures to include reflection and decision-making” (A/NZM4).

However, the process for the development of the “SOI is top-down with key intentions key achievements. Research participants promoted using tailored sets of service indicators at service level” (A/NZM6) and “outcome frameworks allow quick field measuring” (A/NZE8) as “need to know impact so can make decisions prioritisation” (A/NZM14).

Integrating management roles, organisational approach, and locations of functions were highlighted to enhance reporting as “currently as institutions are using a more list approach rather than evaluative” (A/NZE3). Research participants considered agencies were “short on resources and needed a regional director to oversee and identify results gaps” (A/NZE4) which may include a “line of sight, context and needs analysis”

(A/NZM8). This finding correlated with the AR rating of seven out of nine agencies showing only some or no progress *for an Increased use of needs analysis* (Criterion 1).

In addition, in both PNG and A/NZ provinces, there was a “need to measure in the regions so can see variances and use indicators that identify regional priorities” (A/NZM12). Longer-term planning and inclusion of evaluators in A/NZ was emphasised:

If [you] have no longer-term plans, you can have a reactive organisation. The focus on evaluation can depend on the drive from the leadership team, which depends on how much pressure there is to perform. Evaluators have skills that go from top to bottom. (A/NZE2)

Whereas in PNG, using “sector strategies links how it works together” but clarity over which agency was leading appeared required to support development partner coordination (PNGM1).

Looking at the sector strategies, accessibility, important that you know how it works together. And but it’s not clear who’s the lead agency, and for example water and sanitation actually has no sector where that fits in and that’s not clear. So, there’s some work for PNG on who’s leading what sector, how it’s driven and the goals within that and how the donors are aligned in behind that. There’s going to be one aid strategy and the donors come in behind that, which is linked to their national framework. (PNGE1)

### ***Communications and organisational culture***

In PNG “leadership” and “the mindset and capability of staff” (PNGM10) were emphasised as ways to enhance service implementation. An “understanding of why use a results approach” (A/NZM4) for organisational culture was also highlighted in A/NZ as:

this is not just about systems and processes, it’s around a culture and a change of behaviour and the recognition of people seeing the awareness and understanding of the value of what we’re trying to do here. (A/NZE9)

Communication was emphasised as critical by several participants as there are “different ways to communicate with different audiences and cultures” (A/NZM4). It is “important people know what outcomes they are contributing to” (A/NZM8). However, research participants considered it can be:

a kind of complex and slightly fraught conversation because of the relationship, what I would perceive as the degree of people’s understanding about the relationship between outcomes and measures. And specifying outcomes in meaningful ways and then being able to move back from that to think about indicators and performance measures. (A/NZM18)

One specific focus A/NZ agency when involving different communities in regional locations reported enhanced relationships and a sharing organisational culture:

We have a relationships team up here that works with the Pacific communities to provide that voice to our policy team based in Wellington primarily. And we also have a relationships team in Christchurch, so the research team has the sensibility to move across the relationships teams that we had as well as within our policy and research directorate. That's how we're working on our priority areas. (A/NZM11)

However, issues over the sustainability of resource use and achieving sufficient results were illustrated by an A/NZ manager as:

I want to talk about management things as opposed to actually what business are you in, what outcomes? Creating the impacts. So that means things are stopped, and people are going down the road. So, once you've got the output mix, they then stitch the outputs through to your end outcomes. The risk there is you might not have enough outputs to create really enough oomph on your intermediate outcomes to actually achieve the end outcome. (A/NZM2)

A question "over who collates results" (A/NZM9) was also raised.

Nobody has a job here at the Ministry of collating all that information or providing a consolidated view of how we're doing in achieving our outcomes. That doesn't happen anywhere unless you argue that it happens through the Annual Report. So, there's no regular reporting that goes to the Executive for instance that says, across the six health targets we're here, DHB deficits are here, and across these 20 other measures of our outcomes we are looking like this. So, it actually looks like our influence in the sector is going either in the right direction or the wrong direction. (A/NZM9)

Pieterse (2010) suggested using "culture as a device for nation-building" (p. 65) to assist in overcoming colonial legacies. He emphasised that a "national culture will be developed by and emerge from the real people" (p. 65). This approach appears to have worked in PNG where the development and use of the MTDP has brought the different provinces, regions and tribes together to work towards shared development goals. Reflecting on this approach, using a theory-building lens, I wondered if A/NZ could use an inclusive collaborative 'new' national planning approach that could support more inclusion, equity, and sustainable development.

## ***Organisation culture and learning***

The research highlighted that improvements were required in organisational “culture, system capability and sharing to recognise value not just processes as we need to have a system so people can learn” (A/NZE6). My analysis showed there was:

tension between doing evaluation and supporting internal teams. The role of internal teams for evaluation and research is to socialise findings within agencies and you need a commitment from managers and leadership team. Evaluation capacity inside organisations to support managers is under-resourced and planned. (A/NZE7)

My research also showed there was a “need for personal commitment of agency personnel to take this evidence base and evaluative management seriously. It needs tenacity and willingness to get to grips with design data analysis and use” (A/NZE8). Some participants suggested there were challenges created by the “independent units” causing a “limited use of feedback loops and thinking about cost and results within organisations of outcomes, values and contribution of projects. For a continual flow of information learning – we need to set up for it” (A/NZM7). A solution put forward by one A/NZ Manager highlighted:

you can make changes in frameworks by linking both outputs, results and the difference. You actually have to have an organisation wanting to hear bad news, not scared to hear bad news, not scared of learning that maybe things don’t work and then learning from it as a learning organisation. (A/NZM14)

However, the research highlighted tensions where “if you go back to the original post Public Finance Act specification of outputs and outcomes, Ministers would choose outcomes and the public service would choose outputs” (A/NZM3). This can lead to “hiding stuff done and not want to know – need to become more mature in having discussions and looking at other jurisdictions” (A/NZM3). “Using an outcome framework to clarify what is valued for the community, provider, regional, and national, and for outcomes needs indicators and questions so [it] can get measured so you can know how its progressing and you need to show perspectives” (A/NZE9). This approach was viewed as assisting communication internally and externally.

Integrating evaluation was promoted “so it is done as part of operations, so it is not done to you as a constructive. We need to link an evaluative team of experts integrated throughout and through a community of practice integrate expertise into the agency” (A/NZM9). However, the “culture of ministries can be risk-averse so [this] contrasts with getting actual information and adapting sometimes challenging with negative info and

how to respond to and a fear of conflict. People delivering challenging news can get blocked” (A/NZE2).

In addition, the “need to do more than anecdotal info” was highlighted as “until you actually need to get a measure of it and it immediately tells you whether you’re moving to it or not, it is pretty hard to have more than just an anecdotal discussion about things you’re trying to do to achieve that outcome” (A/NZE4). This requires “links between outcomes and learning [with] a systematic approach that is flexible and adaptive. You need flexibility in the system to get innovation and prioritise” (A/NZM8). This could be achieved by “using evaluative questions and narratives to get as robust as you can changes over time through quantitative numbers and qualitative narratives which links policy, planning and evaluation” (A/NZE4). Another consideration was that “learning is happening at the community level – evaluative learning. Evaluation is still being studies and can be unwieldy, [you] need frameworks into community levels so can feed results up for social and economic” [outcomes] (A/NZE11).

However, it appeared functional structures for integrating evaluation may also require explicit reconsideration.

There’s no function in the Ministry that has that as its reason – we’re going to improve the performance of what we do as an organisation. Nobody holds that hat. This requires a culture of reflection, learning and reporting. There’s no systematic gathering of ideas, gathering of information, improving what we do – nobody has that job so in the long run that is where we need to build strength. And our focus heading into this year after crisis management last year has been much more on our internal performance reporting processes. (A/NZM5)

A solution put forward included:

we’re so big we have so much information – it’s actually cutting it all down to that which is an absolutely of value. It’s simplifying what we provide and making sure that it’s giving them what they need because it’s up with the play in terms of what’s changed operationally. (A/NZM9)

Furthermore, the importance of enhancing regional inclusion and participation was emphasised by some participants in both PNG and A/NZ. Having “offices around the country assists for community links” such as Work and Income in A/NZ. Participants considered this approach could be expanded to include “family and community perspectives” (A/NZM12) which were provided by the 22 provincial offices in PNG.

Overall, my findings pointed to an increased use of results-based approaches and partnership relationships, expanded communication and knowledge sharing

organisational cultures. These areas require expanded practice and could be enhanced with increased inclusion of key groups such as providers and communities.

#### **6.3.6. Integrate management administration and process data, monitoring, evaluation, insights, reporting and learning**

In the sixth substantive area of the institution level indicating attention required, this included extended management relationships for sector, organisational performance and increased clarity with their overarching frameworks

there wasn't any sector strategy thing going properly and there wasn't anything around organisational performance – those are other parts of the Ministry which we have discussion now, because we're working quite closely with sector strategy cause it's not just about what the organisational performance directors looks after, how is it that we know if the [Ministry] was doing a good job, or there's this whole thing of how would we know if the sector as a whole was going well. So, you've got

However, within an agency there appeared to be “two levels for evaluation – strategic and activity level or programme level which usually comes through the strategic level evaluation where the activity level is based on the policy. So, if it's got high value, it's deemed strategic, or there's higher risk, they will evaluate” (A/NZM5). Links with internal research and evaluation and cross-agency teams appeared needed. One participant suggested using “strategic frameworks which link teams and connect people as it is important to have discussions” (A/NZE7).

One agency reported expanding guidelines, policies and evaluative capability which appeared to assist as the “new policy and guidelines that we came up with last year did have clearer instructions and expectations around when monitoring occurs and when evaluation occurs” (A/NZE5).

However, it appeared that an increased focus on “learning and accountability” (A/NZM18) was still required to “get more traction for action” (A/NZM8).

There still appears a methodological lack of clarity of integrating M&E within business units at operational levels to enable aggregation of data up to strategic levels and particularly in evaluation land some of those things... it's easier in an evaluation, you can do case studies or you can sample. But for our approach, we need to have rock-solid data. So, it will be meeting our evaluation colleagues halfway but also pushing our business units hard because Audit's [A/NZ] asking – how are you going to prove this? (A/NZM12)

Another consideration highlighted by participants was “at which stage would we be able to expect some of the early impacts to start manifesting and what information will we need to actually report back on that?” (A/NZM5). Moreover, the findings highlighted that linking management administration and M&E data was viewed as important as it links data on people and relationships, which can be enabled by integrating functions. This was emphasised by one A/NZ manager where “links with performance, policy cross-agency, inter-agency and component evaluation teams within a support service area were more aligned with our internal performance management, performance reporting processes and functions, and the research team” (A/NZE10). Alongside this, the value of using qualitative and quantitative data through M&E was emphasised at operational levels.

Some of the agencies believe numbers to be a form of heresy and what they will say what we do is important – you need to understand the human dynamics. But, we believe that sometimes we need to bring our colleagues and Crown entities a little closer to appreciate that numbers ultimately can matter in certain circumstances, as well as us moving it into their territory to understand why some of their qualitative perceptions matter. So, I would argue that monitoring is fundamentally important in forming an effective relationship between people, teams and institutions about these matters. (A/NZE8)

This research noted integrating data collection with collaborative planning, identifying key data and undertaking analysis within business units appeared as an evolving practice which may require further attention to become more systematic within agencies.

#### **6.3.7. Agency planning, reporting and provincial coordination**

In the seventh substantive area of the institutional level indicating attention required, my analysis pointed to the planning cycle for agencies and ministers in A/NZ’s electoral cycle, managerial theory and processes for planning, policy, and evaluation as all appeared to impact on development effectiveness and accountability. It comes through strongly in the interviews that election cycles can affect planning priorities in A/NZ.

The three-year election cycles as well and with a change of governments, you get changes in policies and changes in the sorts of things that they want to achieve, which again has to be reflected back into agencies at the end of the day because an agency is an instrument of the government policy at the end of the day. The government devotes the monies for what the agencies do. (A/NZM22)

Therefore, having agencies and sectors with an institutional view for short- and longer-term planning, implementation and measurement appeared important for enabling



adaptive services and interventions. A balancing issue between agencies guiding ministers and ministerial priorities was highlighted as:

the individual at the head of the Ministry is changing all the time and his [or her] personal views whilst important, and that's what he [she] was elected for and put in and are clearly important. The Ministry view, which is more enduring, particularly about the organisation seems to just evaporate. (A/NZM5)

Research participants considered having a Minister and agency agree on shared priorities was important, and this may involve cross-agency collaboration.

it's challenging because in some ways it's hard for individual Ministries with their own Ministers to work together because they've got their own accountabilities to the Ministers, but there's a bit of a movement towards joint accountabilities. (A/NZM6)

However, with "layers of management, governance, information flows, and recording that sits under results, to distil it can be quite complex" (A/NZM21). This appeared to require "clarity of alignment, funding and contribution story. Some agencies are working hard, and others require more effort and clarity" (A/NZM4) as "values and what is valued needs to be more explicit so you can get clearer focus on priorities" (A/NZE11). An A/NZ manager outlined how:

one of the hardest things around performance measurement is accepting that it is a values-based business thing and usually the values aren't spoken about explicitly in these environments, but quite often it's sort of like they can be seen as almost like some common-sense. (A/NZM7)

Furthermore, this research showed that regular inclusion of key groups in planning, analysis and reporting was not explicitly outlined or occurring systematically

So how does the Ministry's thinking about outcomes actually encapsulate all of that in a balanced way? And to some degree, we have to reflect the values of the government of the day, but we also need to inform them in an ongoing way as well. (A/NZM5)

These findings raised concerns over the nature of democracy where Ministers are elected to represent the views of citizens. Davies (2007) outlined that a more "deliberative democracy" was required which "demands not only participation but also equal access to decisions by citizens with a stake in them" (p. 81). My findings concur with White's (2009) views as A/NZ's devolved public sector managerial approach does not appear to ensure systematic inclusion or participation and that more deliberative centralised approaches may be required. PNG's mandated centralised national planning

processes enable more systematic inclusion and provide greater transparency in national planning goals for its citizens.

In terms of wider accountability and impacts when using more sector-based approaches the effect was widely seen as positive:

it's been excellent, and it's one of the few ways in which ...a Ministry can actually significantly influence the behaviour of grassroots organisations and health workers. The surprising thing is that it's happened purely through reputational effects, so there's no financial reward or punishment for achievement. There's no formal sanction of that kind, it's purely the fact that it gets published in the newspapers is very effective in improving performance. There is now a consensus that in the A/NZ context it's worked very well, and suggests it will be a way of working with the sector that will only grow, at least in the medium-term future, as a way of setting aspirations and goals and tracking whether we're moving in that direction. (A/NZM9)

Likewise, in PNG, there was "growing interest in reporting sharing, starting to see the change PNG wants to have. The development results at country-level based on the international Paris Declaration are guiding to show how PNG is doing – including women and children, so we can demonstrate for every sector in PNG – what is being spent what is happening" (PNGM4). However, "[we] need to commit funds to it. For the OECD-DAC, PNG signed in and to use a reporting system that should be based at central agencies. National Planning (DNPM) holds the development budget for the country – health, education – making sure it is invested and accounted for (A/NZM10).

In addition, increased "capability in the wider government sector" (PNGM9) appeared required as there was "limited understanding of outputs and outcomes. The National Planning Department (DNPM) has limited capability and understanding, which in turn is limited by capacity and knowledge. It's everybody's business to do M&E – people need to know how to do it and how to benefit from it" (PNGM5).

Moreover, provincial planning and coordination were promoted as important and the positive use by PNG of coordination committees was highlighted which cascaded down from national to provincial to local levels.

A few of the provinces have committees as the inclusion of the provinces is important and it enables tailoring of priorities for provinces. The provincial coordinating monitoring community is led by the provincial administrator. And all members are all national agencies or state who are down on the province positions like Telecom, power and water board. All come together – police, agencies come together including civil society and private sector, the Chamber

of Commerce, for instance. They're also part of that. So, we have seen a lot of difference since these committees, and I think now we have covered almost all the provinces. (PNGM10)

This increased coordination appeared to be resulting in expansion of services with one PNG manager providing an example:

We have really seen a lot of improvement, an example was in the fire service – they are located in a very small area and they couldn't expand because the province is growing ...the city at the time was growing so they needed new land. Now they have been given the new land for that. And in the past that would not have happened. The provincial administration operating is ongoing and so are those are coordination committees. (PNGM8)

Moreover, the role of monitoring and reporting was emphasised as important in both PNG and A/NZ regional contexts as “we want to help more other regions to be able to do, have the skills of monitoring, and that will really help us to improve. Because I believe monitoring is a critical part” (PNGM8). An A/NZ manager highlighted why data collection from regions might require a systematic coordinated approach with inputs from agencies as:

in the current context, we need to be really mindful of who are the people who would be collecting that data. Our people in our regional offices, we need to have an awareness of what those people do and what would their capability be like for collecting different kinds of information, for using different kinds of software? Whereas actually it's the agencies that you have to deal with that data in the longer term. (A/NZE11)

Overall, in PNG and A/NZ, research participants indicated some improvements. Yet, in PNG, the flow-on effect to improved development effectiveness was mixed with “people feeling lives are improving, but there's a data gap showing as it is not evident” (PNGM5) and “there is room for improvement in results from programmes and projects” (PNGM8). These findings were also reflected in A/NZ context, where “reporting and transparency” were noted as “occurring and improving” (A/NZM10), yet further attention was required “as there is a need to do more” (A/NZM3).

My findings showed that dimensions included in the updated impact model at the *institution* level such as *Improved management processes* and *Improved M&E processes* required further attention in both PNG and A/NZ as interviewees noted there were insufficient levels of relevant data available particularly on outcomes. This outcomes-based data is fundamental when using evidence to assess development effectiveness and in decision-making.

## 6.4 Emergent concepts requiring inclusion

The following section outlines emergent research findings associated with the *institution* layer. The key findings were highlighted in Table 9 and was followed by a narrative.

**Table 9. Emergent concepts requiring inclusion: institution layer**

Concepts		Evidence
1	Increased community inclusion in service delivery needs analysis.	Evidenced in qualitative data and AR. This includes participatory approaches, results measurement, and supporting regionalisation with communities of practice.
2	IT systems enable online processes and increased timeliness for data.	This was evidenced in qualitative data, AR and RT. PNG and A/NZ contexts displayed aspects of linked regional hubs for data collection occurring which could be made more explicit and evidenced through use of shared IT platforms.

### 6.4.1. Increased community inclusion

The findings indicated that increased community inclusion in service delivery needs analysis, participatory approaches, results measurement, and supporting regional coordination may enhance development effectiveness and resource allocation. This was evidenced by the qualitative findings (QF) and accountability document review (AR). A narrative of key findings followed.

The first emergent finding for the *institution* layer highlighted a basis for “rethinking evaluation as a discipline” that operates as a “corpus of work” such as national, institutional, regional, thematic, and outcome levels rather than being “programme specific” (A/NZE3).

I think people tend, have tended to think about evaluation as being programme specific as opposed to a discipline which can be applied to a corpus of work which is related to one broad outcome. New pieces of work, including the redesign of how professional development is allocated to schools, new contracting arrangements for professional development the way it is. There are changes in the regional functions in the way they work with schools. I think we’re starting to get some traction around framing evaluative capability as a requirement in terms of using information, to adjust what you’re doing as you go along. (A/NZE3)

Taking a systematic approach for M&E and linking to business plans was emphasised by research participants. A systematic approach could involve “looking at information needs and coordinating to streamline activities and integrate with provider reporting and

increase their evaluative capability” (A/NZE10). A solution put forward by one A/NZ manager outlined using a more participatory approach which included:

...recently embedded explicitly in new contract monitoring and reporting processes. We’re asking providers to use, to report around a very simple intervention logic and so we’re asking them to assess quality of their provision around this range of domains, and they have to demonstrate to us that they’re performing at the level that whatever they think they’re performing. (A/NZE3)

Building evaluative capability was also advocated through organisations, providers, NGOs and community groups:

There’s a need also to build capability through the organisation, and with policies and guidelines, we can structure good training to support the organisation. Previously, it was a little bit difficult to provide that training because things were a little vague on what the expectations were. Now, we are able to all preach from the same policy and guidelines which is better for the organisation. (PNGDP)

My research findings pointed to recognising there are “different realities for different communities” (A/NZM6) and “different perspectives” (A/NZE8) in both PNG and A/NZ contexts. A question raised by research participants in terms of community inclusion was “who is defining outcomes at the community level, and at the national level?” This was because “I think potentially for any outcome, there’s always the ways of measuring it and different perspectives on the outcomes as well. It’s kind of the whole multiple realities” (A/NZM8). Taking a sector approach appeared to support getting “agreement across different agencies” (A/NZM13) and perspectives.

Whereas in PNG, inclusive approaches also needed to involve the “informal sector” (for example people selling produce at markets) and aspects such as “sustainability of programmes” and in “regions and communities” as PNG’s “informal sector policy encourages people to be independent. Community development, once given opportunity are strong, resulting in the sustainability of programmes. There are community workers providing extension support out in the regions and communities” (PNGM9).

In addition, in PNG, how “beneficiaries/people can access service support in the regions which enables people to get [the] bigger picture. PNG nationally has strong provincial and monitoring linking service delivery” (PNGM6). Interestingly, the terms “decentralisation and regional effectiveness” were promoted as useful emergent concepts by research participants to assess “follow through with what is happening on the ground as provinces manage their own services’ service delivery and there are challenges in different areas. It’s good to discuss with regions to support and find tailored solutions for provinces given geography and people” (PNGM8).

The findings suggested that agencies may “need frameworks into community levels as evaluation is still being undertaken as studies which can be unwieldy, so you need frameworks into community levels so you can feed results up for social, economic, and wellbeing. If you have evaluative discussions and thinking, then you get more traction for action” (A/NZE8). These views aligned with Innes and Booher (2018) who advocate for “genuine dialogue” (p. 18) and participation of communities in their development. As an evaluation practitioner, I considered this be a critical way of enabling more participatory, and adaptive responses evidenced through use of shared platforms.

#### **6.4.2. Information technology systems**

A second emergent finding was that IT systems appeared to be enabling emergent and innovative online processes and increased timeliness for data, as “going online will speed up processes” (A/NZM4).

One of our bigger projects that links really strongly to this is if we can get all that online, it will speed up a lot of processes which will improve linking processes and sharing. (A/NZM8)

The “value of having national standards and sector health data-sharing” (A/NZM10) was emphasised as it enabled analysis and focus on certain areas such as primary health care. However, “one of the harder things with information systems is to keep them evolving and keeping the standards consistent if you’re looking at something on a national level for any good standards you need good systems for collecting information” (A/NZM5).

Therefore, it appeared agencies “need systems for information linking info needs and IT systems to assess effectiveness” (A/NZE9). An example provided by one A/NZ manager in linking IT and strategy included:

What we’re doing now at the moment with the IT strategy and the big health idea strategy is development of an electronic health record, moving more spending into primary care it strategically makes sense but there was actually insufficient attention given to collecting the information we need to understand how effective we’ve been. (A/NZM15)

Research participants advocated for more regional data collection as “being able to get data is important particularly for regions so you can analyse locally. We need quality in and more skills of people with data” (A/NZM7). Retrieving data was also highlighted as an issue which appeared to require further attention.

We’re actually in the process of being able to do that ourselves we’re learning SaaS at the moment. This allows regions to receive the information faster and

they can access it and interpret it and play with it themselves which is good.  
(A/NZM5)

In addition, sharing of data through IT systems assisted “getting people on the same page and measures and impacts are key with communication, performance frameworks and IT systems” (A/NZM6). However:

Sharing of information is something from an IT perspective that we’re still finding difficult because we have evolved from often quite separate businesses which weren’t even with us. So, we’ve developed different IT systems using, putting data into different places and sharing of that information has been problematic because we can be on page one, but IT systems take a very long time to be able to evolve. (A/NZM6)

Another key issue for improving development results highlighted by research participants was being able to “aggregate data” (A/NZE8). But this involved overcoming “the challenge of getting the right measures and impacts that are actually the meaningful ones and not just the ones that are ones that are probably palatable to everyone. Let’s moderate them down so that’s manageable” (A/NZM4).

In addition, my research also noted implications for data standards and security, meeting local needs, integrating planning frameworks and reporting, data sharing, sustainability of service delivery and data collection, and participation of providers and communities in data collection and analysis as emergent areas for consideration in terms of enhancing development effectiveness and governance.

#### **6.4.3. Summary**

The findings reported in this section pointed to the need for achievement of transformative change from agencies and sector required more inclusive knowledge sharing to adapt and enhance services and interventions for people. An emergence of thematic groupings such as community development in agencies appeared to be assisting agility and learning. These findings raised questions in the A/NZ context over the focus within central agencies on CE’s rather than senior management and regional leadership teams. This research suggested that with further flexibility and integration rather than restructuring within agencies, development results and effectiveness may be enhanced.

The research suggested that A/NZ’s limited country approaches to development and evaluation impacted on inclusion, equity of service delivery and interventions, and systematic ethnic, regional and community participation. Whereas, in PNG, the country plan provided direction but how it was implemented, and progress measured required

further implementation design and monitoring of ongoing accountability to ensure services and interventions were reaching citizens and were sustainable.

## **6.5 Chapter summary**

This research pointed to an expanded role and follow-through from Audit NZ that is improving agency focus on outcomes and performance, increasing transparency and accountability. In PNG, auditing for financial and non-financial information appeared very limited, and participants believed it required a significant focus, political will and government commitment to improve accountability at all levels of government.

The findings also highlighted that more systematic community inclusion and participation can be improved by providers and communities being more involved and through 'real time' data collection which enables ongoing feedback on services and interventions. These approaches aligned with using a more explicit values-based management approach and promoted increased inclusion, partnership and participation. In turn, findings suggested these approaches may enable more responsive, iterative and adaptive evaluative inputs, customer and programme focused IT design, and improved communication and learning culture.

Research participants considered these integrated and adaptive approaches were potentially more transformative than the transactional management processes which evolved from the late 1990s public sector reforms. Improved interpersonal relations and adaptation by managers appeared required in collaboration with providers and communities, supported by expanded evaluative and business model capability and integrated IT systems.

Areas considered important by research participants included increased strategic outcomes leadership; integration of development, evaluation and management through a state sector system business model; a more explicit function for evaluation; enhanced central agency role with the use of results and outcomes frameworks; alignment of development partners (national, NGO, private sector, donors) to country priorities, and increased accountability.

Chapter Seven follows outlining findings for the *wider infrastructure* layer. A critical discussion of the findings was undertaken in Chapter Eight.



## Chapter 7: Wider infrastructure layer

The third layer of my analysis focused on the *wider infrastructure* findings and addressed sub-research question one: What evidence in A/NZ and PNG demonstrates what works for whom in relation to results frameworks and associated management systems? The *wider infrastructure* layer specifically focused on the role of central agencies, their guidance and interface with line agencies.

### 7.1 Rating of impact model dimensions – wider infrastructure layer

The Rating of Impact Model Dimensions undertaken by research participants associated with the *wider infrastructure* layer during the A/NZ case study was provided in Table 10. This table mapped three improvement dimensions against four simple standards that ranged from “Consistent progress evident” through to “Unsure” (whether improvement is evident or not). These ratings were based on good practice concepts identified through the literature review. The table synthesised a range of findings from the data sources described above and represented the researcher’s subjective view of what was reported across those findings. This rating was not undertaken for PNG due to the more limited level of case research.

**Table 10. Rating table for impact model dimensions: wider infrastructure layer, A/NZ case study results**

Impact model dimensions for <i>wider infrastructure</i> layer (taken from Figure 9). Research undertaken 2014 (N=9).	Consistent progress evident	Some progress evident	No progress or change evident	Unsure
1. Increased development effectiveness	2	1	3	3
2. Improved aid effectiveness	(not relevant for A/NZ case study)			
3. Improved lives for people and environment	1	5	1	2

Two observations derived from this rating table for the *wider infrastructure* layer can be related back to the core criteria in Table 1.

First, while one-third of research participants rated some or consistent progress is evident for *Increased development effectiveness* (Criterion 1), six out of nine of the participants (N=9) rated either no progress or unsure. This latter finding can be tied into the three out of six agencies who rated poor for *Increased use of baseline data* (Criterion 5, Table 1 in the AR) on the assumption that if baseline data was not well established or

was poorly used, an agency was unlikely to be confident of where and in what ways it measured success.

Second, the above rating can also be read alongside the *Increased use of evidence-based data* (Criterion 1, Table 1) which was rated poor by three out nine A/NZ participants and also rated poor in terms of the AR rating. To be consistent in *Increased Development Effectiveness* (Criterion 1) an agency required robust evidenced-based data, which also extended to an agency's ability to report confidently on either assessing sustainability (OECD, 2012) of development results or measuring the progress made towards *Improved lives for people and the environment* (Criterion 3).

## 7.2 Substantive progress – Wider infrastructure layer

While the RT provided a snapshot of improvements in key dimensions identified from the OECD-DAC guidance literature (2005b), which was the entry point for this research, this next section focused on what was described as the substantive conceptual findings in both A/NZ and PNG. That is, as discussed in Chapters Five and Six, these were findings that pointed to knowledge and practice ideas at a conceptual level, organised to assemble, on the one hand, those ideas that suggested that progress is well underway and on the other, those that required further attention or are poorly conceptualised at present. The full table is provided in Appendix F, while one excerpt highlighted here (Table 11) provides the substantive conceptual findings showing progress for the *wider infrastructure* layer.

**Table 11. Substantive progress: wider infrastructure layer**

Concepts		Evidence
1	In A/NZ, expanded audit guidance and standards for non-financial information and increased use of outcomes/result frameworks by agencies	Evidenced in accountability document review (AR), rating table (RT) and qualitative findings (QF). An increased guidance and audit focus in A/NZ was assisting to clarify development goals and outcomes.
2	Increased use of A/NZ agency and sector outcome/results frameworks.	Evidenced by AR, RT and QF. This provided published direction for sector and agency strategic intent and funding alignment.
3	PNG MTDP provides published direction for PNG country and sector strategic direction budget allocation	Evidenced by AR and QF. The MTDP 1 and 2 has supported alignment of development partners. However, implementation challenges in regional service delivery incurred by agencies in PNG impacts service coverage, funding flows and results-based data collection.

Concepts		Evidence
4	PNG planning framework mandate and provincial and local-level governments	Evidenced by AR and QF. The PNG planning framework and Organic Law supports systematic provincial inclusion in planning, coordination, administration, and service delivery.
5	Positive aspects in PNG and A/NZ in regional service delivery coordination centres in sectors such as law and justice, health, and transport.	These regional centres balance devolution and centralisation enabling some adaptive inputs for different contexts and needs of regions and communities with national coordination and support. Evidenced by QF, AR, and RT.

### 7.2.1. Expanded audit guidance and standards for non-financial information

The research highlighted the value of the proactive role of the central agency Audit A/NZ in seeking to improve performance in the A/NZ state sector, and the observed improvements. Improvements noted in government audit reports were accepted as showing that:

those results are very, very encouraging – it has got a four-point scale, so it goes from ‘poor’, ‘needs improvement’, ‘good’ and ‘very good’ – and the assessments in the scale and the scope is included in the B29 to Parliament. You’ll see it kind of trending upward. When we started on this journey, a few entities were ‘poor’, and then ‘needs improvement’. There are far more of them in the ‘good’ category now. (A/NZM1)

Agencies and Audit NZ were “motivated” (A/NZM20) and provided improved high-level guidance (A/NZM3).

What has also helped is organisations have been motivated. Auditor General standards have changed, but the office has also issued some better practice guidance – on the OAG website – and there’s a section on service performance with all their reports that some of their more recent reports are looking at entities and looking at their performance frameworks and some of the better practices so other entities can learn from them. This falls broadly within the bailiwick of issues of importance to the Auditor General, the performance management and information about the results of public entities and results across government, and enduring concern. (A/NZM20)

However, one A/NZ manager considered “there has been a longer-term issue of accountability and a question over how much to guide and how much to devolve to agencies” (A/NZM18).

I think is that there is a fairly widespread perception that there is somehow a statutory failure and I think the statutory requirements are actually quite small and quite flexible, but nonetheless, the embedded practice has interpreted the statute in a bunch of ways. And that is creating a drive for more flexibility. (A/NZM21)

Therefore:

as long as you have a system that provides the answer to those basic couple of questions – Am I delivering my services effectively? Are they having the results that they are expected to have? And have I got the capability in the long term to keep doing what I am supposed to do? (A/NZM21)

However, my analysis showed there were opposing views between line and central agencies on performance management and audit processes. Expanded guidance was wanted by line agencies that were more prescriptive with more directive support from central agencies (A/NZM6). Whereas central agencies considered:

Where things will go is to say less emphasis on the prescription of what you had to provide and more emphasis on the purpose of what it's for. But the challenge in that will be, as has been the challenge in the last 30 years that we will need more intelligent discussion and implementation of that because I think there has been some doctrinaire and fairly coarse and simplistic thinking. There's also been some perennial neglect across organisations within and the public sector broadly. (A/NZM5)

Both central agencies and agencies managers concurred that:

We will have to look ourselves in the eye a bit more and start saying, what is the value we anticipate? Why are we putting this investment in, and what do we expect? more than, things just are because they are, and that will be uncomfortable both for managers and frankly at the political level if that occurs, it's inevitable. (A/NZM15)

Overall, there still appeared to be a gap between central agencies providing sufficiently detailed guidance in A/NZ for non-financial information which central agencies consider as operational, and agencies wanting more detailed guidance to assist and standardise practice. However, PNG's 2016 mandate for evaluation was for this to be coordinated and undertaken alongside the MTDP, placing strategic evaluation at a country level. This evaluation mandate in PNG follows an emerging global trend in

countries such as UK, USA, Canada, Ghana, Chile and South Africa (Bamberger, 2010) which are examples of an increasing drive by governments to be able to assess and report on their development progress, which appears in line with a focus on sustainable development (UN, 2015) and adapting to issues of climate change and growing inequalities.

### **7.2.2. Increased use of A/NZ agency and sector frameworks provides direction and funding alignment**

In the A/NZ context, my analysis pointed to increased use of agency and sector outcome/results frameworks was evidenced by the agencies' strategic intent and funding alignment published in their SOIs. In addition, there appeared to be an evolving sector alignment within the justice and social development sectors which were more coordinated than other sectors. Two issues highlighted included a sense of A/NZ agencies being "more compliance-driven at the beginning rather than strategy-driven" (A/NZM3), and because agencies "know about high-level outcomes, spend and keeping admin records" (A/NZE3). Both were demanded through audit processes but were not necessarily viewed as useful in building capacity in agencies to transform practice. An A/NZ manager outlined a reason for this gap as:

there's too much built into the transitory, transactional information and not enough 'what are we using it to try and understand them to do'? So, I think there's a whole demand-side problem about if we had all this information what might we do with it, and why is there a demand-side problem? (A/NZM15)

My findings indicated the use of outcomes frameworks and intervention logics was assisting with transparency and accountability at output level (A/NZM5). Yet another manager questioned the focus on agency outputs rather than outcomes, and that lack of awareness about "whose reality is it?" (A/NZM8) that is affected by demand-side compliance. A further issue raised by participants related directly to performance management and results frameworks:

Agencies are being encouraged – because there's a big push for managing for outcomes – they're being encouraged to sort of start with the outcomes and then work back to what they're doing, whereas I think agencies need to be thinking less about just meeting – complying with the system – and more about actually making performance management work. (A/NZM14)

A second manager added that "there's nothing that's just a performance issue that isn't also a financial issue" (A/NZM20) implying perhaps that more attention needed to be paid to the connection between financial and performance reporting. Moreover, the

A/NZ participants understood the politicised nature of agency performance frameworks. As one manager stated:

because of the nature of – and also the fact of life with politics – you have to start with the goals that the government want to set in place and then frantically try to work out your intervention logic between the stuff you can control and how on earth are you going to have any impact on that. (A/NZM19)

Another suggested managerially, “that could be very difficult then to draw those clear lines of accountability, if you don’t have a clear understanding of how your activities link to your outcomes” (A/NZM8). As one participant queried, where was the focus “on what matters?” (A/NZM10). This question was amplified by another manager as:

... what have you actually put it at the heart of the thing? No one’s at a seminar or workshop where Treasury are looking at the next steps of a public management system. And I said, well, you’ve got to make your mind up, is it inputs, what is [the] basis? Is it outputs, which is AG4, or is it outcomes, which is managing from outcomes? What’s the premium mobile – what drives the system – what drives the expectations on people – what drives performance assessment? They didn’t sheet it home and say, “This is what I review: chief executive performance on X. The chief executive was an outcome manager, and that’s what they were held accountable for. They sure would make everybody underneath them accountable for it. (A/NZM12)

Finally, a sector stakeholder reflected that:

... benchmarking – looking at practice in A/NZ and the outcomes – the practice in A/NZ is interesting. It’s quite variable – in terms alignment and a line of sight, or activities that are contributing to outcomes, or do they know what direction they’re heading in – that doesn’t appear always to be at all how it’s working through as people quite often substitute these [outcomes] kind of for process and output objectives. (A/NZM16)

Overall, the QF brought a range of issues in to focus including that, on the one hand, the use of frameworks was increasing, and on the other, that there was confusion about what the themes underpinning those frameworks should be. It appeared that increased central agency leadership and guidance might be required to enhance progress towards sector and country priorities, and increased accountability. These findings were supported by more recent calls by the OECD (2019) for new approaches that include “vision, evidence and capacities at all levels of government” (p. 37). This led to further consideration that the paradigms of performance management that are undertaken

without using qualitative data from key stakeholders may be constraining more inclusive and equitable approaches.

### **7.2.3. PNG plan enables country and sector direction, portfolio-based funding and alignment of development partners**

PNG's country plan (MTDP) provided direction for country and sector strategy and budget allocation and supported alignment with development partners. Participants noted that PNG's formalised planning approaches highlighted more explicit development planning and prioritisation. For example:

the mid-term development plan (MTDP) basically is aligned with all the sector plans, and also with Treasury and the funding, and then linked to recurrent budget and development budget which is linked there. This [plan] builds on the midterm development strategy from 2006 to 2010, and out of that came seven enabling priorities. (PNGM5)

In addition, a senior manager outlined how the planning approach was used as a collaborative tool by DNPM, which also included donor alignment.

It basically, involved looking at high impacts, customised, tailored to meet the local context, gender, and working together, how to meet the mid-term development goals, for 2015 reporting and on both [plans] from 2006 to 2010. This included looking at customised targets with donors to work within. (PNGM5)

However, PNG research participants considered implementation challenges incurred by line agencies impacted on regional service delivery coverage, funding flows and results-based data collection as:

a big challenge for PNG is the geographic and services in the provinces – how you work with provinces down to the districts, the service centres. The remoteness – the feeling was that development partners only concentrate in some provinces and development effectiveness is for all people. (PNGM5)

In addition, efforts appeared underway to enhance data collection in PNG as “there is an absence of an M&E framework which brings together the results from the different sectors. Currently that is the key focus and they're thinking through how to do it and what's required” (PNG M 6). Another area of focus was on the:

role of provinces and how does funding go straight down into sectors and that the current thinking appears that the provinces then become more of the monitoring role – for monitoring the progress in the different sectors – which is an interesting concept. (PNGM9)

Moreover, the sector approach (such as health, transport and justice) used in PNG “relies on partner countries aligning to sector framework and agreeing on inputs” (PNGM7). For GoPNG, MTDP alignment included NGOs, private sector and donors. My analysis indicated that the use of results frameworks in PNG also involved donor country posts (for example, MFAT or DFAT High Commissions located in PNG), which highlighted priorities of partner countries such as community policing by A/NZ. The MTDP in PNG was also assisting countries (such as A/NZ and Australia) to align their inputs to PNG country goals. Although development partners such as this participant expressed concern about what they know and do not know if that advice was not forthcoming.

We know that there are a range of things that affect whether services actually get delivered and address what we think are the key blockages in this sort of delivery chain. But how do we know? What results do we pick out to know whether or not this is actually working? (PNGDP)

However, the alignment of development partners for measurement still required explicit negotiation. A solution put forward by one development partner to enhance “performance” was that donor programmes could align their frameworks against country and frameworks with the partner government

... the sector ones [goals] are much more about being able to measure performance against a whole range of objectives. ... you would need to look at a whole range of indicators ... in terms of being able to make a judgement on whether you’re reaching a particular objective. (PNGE2)

Provincial priority setting and monitoring and reporting appeared included in the planning framework in PNG and supported by the provincial authority agency (DPLGA). One PNG manager said, “the Organic Law Section 119 requires every provincial government to report on their performance and so it’s a mandated requirement, but, since the organic law came in they have not reported because there was no template for reporting” (PNGM10). Coordination issues were highlighted for services in the provinces:

Because we have national governments going down, direct to the provinces. And then we expect provincials to come up [to] the national government – so they report direct to the education department, for instance, health, to the health department. So, there is no coordination ... the reports are going in any direction, and there is no coordination. (PNGM9)

However, improvements were noted by two PNG research participants as:



in the organic law there is a provincial, local-level government services' monitoring authority. [It] is the monitoring authority which is coordinating – shared by the sector or department. It's automatically the chairman of that authority ... who are mandated. And so, he chairs that authority and so have all other agencies, sectors of other agencies they are part of that, and it gives them help – finance, Treasury, planning, other national agencies. Service agencies, they form the part of that subcommittee, and they are required to meet once every quarter, so, as a requirement that they meet and they have their agendas to discuss some of those issues, coordinated issues. (PNGM10)

Yet, participants emphasised that measuring outcomes or impacts at a provincial level still require further attention.

So ...just measuring... only the expenditure but there is not a mechanism that find out whether the outcome is there or the impact. So that is the one area that provinces are also saying that – you are measuring our expenditures, but you are not going down to see what are the actuals? Are they there or not? And whether some provinces are doing the right thing in their expenditure that is not done. This issue [of financial accountability] is for us to really work towards. (PNGM9)

In summary, the QF themes identified the importance of the mandate for provincial inclusion in strategy and planning. Yet there was variability in how this is understood and functions in practice, and the need for further development.

#### **7.2.4. Regional service delivery coordination centres with sectors in PNG and A/NZ**

The use of regional service centres appeared to balance devolution and centralisation in both PNG and A/NZ by enabling some adaptive inputs for different contexts and needs of regions and communities with national coordination and support. In A/NZ, these centres were either in place, or emerging in sectors (such as justice, health, social development and transport) (A/NZM7, 9 &14 and A/NZE6&7). However, an issue highlighted from this research was the varying administrative boundaries which appeared to hinder alignment and coordination across sectors at regional levels (A/NZM14). Similarly, in PNG, regional sector service delivery and coordination hubs were increasingly being used (PNGM7, 8, 9 &10) to assist co-ordination for service delivery in the provinces. A rationale for these service hubs in PNG included:

They are working as a team. First a service agent in the province, they are not working as individuals and the police are working together with the administration, the whole service delivery. (PNGM9)

In addition, the sector approach used in PNG aligned to the country plan (MTDP) was designed to include NGOs, private sector and donors within specific provinces.

There's a sector plan, the sectors are responsible for this and national planning is really using is sectors but also, there's an agreement with NGOs for national planning about providing services, working. Also use the private sector in their relationship, so they signed agreements which goes back to Acra 2008, the churches, civil society 2009. So now churches are really helping support meet the MDG targets. (PNGM5)

The findings highlighted the value and use of having a country plan with sectors as it supported PNG's central planning agency to coordinate and align inputs from PNG's multiple development partners towards sector and country priorities.

### 7.2.5. Summary

This research showed that a coordinated country and sector approach and systematic regional inclusion in PNG and in some A/NZ sectors appeared to contribute to improving development effectiveness with goal prioritisation and enhancing service delivery coordination. However, increased attention to expanding these approaches was called for by both PNG and A/NZ participants.

These approaches such as mandated country-wide inclusive regional approaches and an expanded role for central agencies in development planning, portfolio-based funding, learning and audit could form a basis for a model for expansion as these approaches pointed to enhancing development effectiveness and governance.

## 7.3 Substantive concepts requiring more attention

There were three areas where current practice requires more attention for the *wider infrastructure* layer (Table 12). Each of these were explained in detail following the summary table.

**Table 12. Substantive conceptual areas requiring increased attention: wider infrastructure layer**

Concepts		Evidence
1	Increased strategic leadership, central agency guidance and country development planning.	Evidenced by qualitative data, RT, and AR. Findings emphasised a need for centralised approaches to be underpinned by values (such as inclusion, and partnerships) supports more inclusive, participatory, and systematic approaches

Concepts		Evidence
2	Increased integration of government role and functions and government to government cooperation is required.	Evidenced by qualitative data, RT, and AR. This integration includes partnerships and sector systems approaches, resourcing, non-financial information, alignment of investment, regionalisation of development planning, embedded theory based evaluative approaches with community involvement and feedback loops.
3	Increased focus on transformative development results is needed.	Evidenced by qualitative data, RT, and AR. This includes increased inclusion and participation of development partners with results frameworks and integrated systematic evaluative and flexible approaches, expanded M&E focus as learning organisations, alignment to country systems and inputs and stakeholder coproduction.

### **7.3.1. Increased strategic leadership, central agency guidance and country development planning underpinned by values**

This research examined two different approaches to country development, namely PNG, which used a centralised mandated planning approach and country plan (MTDP). This plan was divided into sectors with coordination cascading from central to line agencies to provincial administrations who appeared to have some implementation responsibilities to support coordination and services down to local level government and ward levels.

Whereas A/NZ devolved managerial responsibilities to chief executives of line agencies and A/NZ central agencies appeared to undertake a predominantly transactional role. However, there were increased calls by managers and evaluators in both PNG and A/NZ to consider evolving approaches to balance centralised development planning with increased regional coordination and implementation. These findings were evidenced by the QF, RT and AR.

#### ***Development principles and approaches***

A/NZ research participants advocated for increased partnership and participatory approaches where such approaches operate as principles, yet this research pointed to mandated approaches being required to ensure inclusion occurs systematically. This research highlighted that in the A/NZ state sector context leaving inclusion and participation up to individual agencies, showed that this does not occur systematically, authentically and democratically for different groups and regions in the A/NZ state sector.

These issues of inclusion and participation were also evidenced in some PNG government and development partner contexts.

It's easier said than done for many reasons. We often are in a rush to design and get some of these projects out the doors, and participation is sometimes impacted upon. Capacity development – that's internally and externally as well. So, we really do want to include our partners in this sort of approach as well, so they're on board. Ownership – if partners don't buy in to what we're doing the chance of success are reduced. And all those Paris principles, ownership, alignment, harmonisation, we're trying to do a lot more work in there, particularly working with [other donors] to get a more aligned approach. (PNGDP)

A PNG based development partner acknowledged that:

The Paris Declaration put obligations on both parties, both on the recipients and on the donors and people as well. One of the ways of kind of resolving these differences is supposed to be the kind of sector-wide approach. So basically the donors would get in behind whatever kind of development objectives the country had for its sector – health or education... and the implication of that was that they would turn their log frames to be consistent with the objectives of the government sector itself, in those sectors, so that you would have the sector-wide approach and both the bilateral and the multilaterals would kind of get in behind that basically and have a consistent set of objectives which was going to be set largely by the government. (PNGDP)

However, variable inclusion and participation in development activities particularly have led to calls for mandating more inclusive approaches within sectors and country planning and across funders (A/NZM16). Participants suggested this finding primarily applied to A/NZ as there was no mandated planning approach for the systematic inclusion of priority groups, regions and communities. These research findings align with Stickl, Haugen and Chouinard (2018) findings on relational power which emphasised that “cultural responsiveness requires active stakeholder engagement and participation” (cited in Chouinard & Cousins, 2009, p. 383). In PNG on the other hand, with the published country plan and mandated planning framework, there appeared to be an understanding that more inclusive practice was required by stakeholders which could include use of evaluation frameworks.

The Paris principles should in fact enforce some harmonisation on the process which should translate through into similar frameworks for evaluation. If, ultimately, everybody is agreed that the objective here is to try and get an improvement in the coverage of clinics in PNG with an end to reducing its

infectious diseases, maternal death rates, which in terms of the detail of log frames for projects, everything would kind of flow from that. Is it working? (PNGDP)

However, research participants emphasised that capacity issues existed as:

in PNG urban centres only covers about 35% of the population. And in most of the countryside, once you get below a small and highly competent capacity people based in the capital, usually the quality rapidly falls away, the further down you get essentially – so that's a problem. (Key informant)

My analysis suggested expanding “evaluation capacity and capability in regions” (A/NZM16) could include “growing local accountability of villagers and officials” (PNGM6). This concept of growing local capability was raised in both PNG and A/NZ contexts where community inputs and participation emphasised: “how we want to improve our collective lives arguing for a very participatory view of democracy” (A/NZM18). Yet,

if you want to have an effective system of evaluation, you have to have talented and committed and motivated people on the ground to make it work? Basically, they have to be committed, to be thinking in terms of the evaluation that you want to make, and that means educated, trained and motivated managers at local levels. (A/NZM16)

In addition, one issue highlighted by an A/NZ manager on the role of government agencies and their officials was that:

the more value-laden ideas of outcomes had less impact on actual practice within management in government departments [in A/NZ]. And, I had disagreements about the extent to which it was possible in fact to manage for outcomes in terms of the ideas introduced really in the Clark administration in the early 2000s. (A/NZM16)

One agency that A/NZ participants considered where some instances of community inclusion have worked is at the Department of Conservation (DOC).

It has a strong scientific community dedicated to issues of local ecologies certainly has this constant idea that DOC workers on the ground are continually assessing and evaluating the effects of their actions to protect the environment in parks and so on, on the actual environment. And they can set themselves some standards of achievement for this and then it kind of continues to be monitored, the quality of the whole ecology that we deal with basically. (A/NZM6)

An A/NZ manager considered approaches could have “dual agency or initiative outcomes” (A/NZM20) as “it’s challenging, there’s a lot of interest in outcomes and government performance, and I think the main thrust of it has not been looking at the dual agency or initiative outcomes but [also] what they mean in total from a government perspective” (A/NZM20). A/NZ priorities were also questioned in the analysis on “what are the current themes? It’s about what does the work of justice, health, MSD achieve in total rather than this is education, that’s health etc.” (A/NZM20). These reflections by participants led to questions on “what the state sector management model means is this annual parliamentary accountability is really the heart of it, but you’ve got all these different players with their own deadlines, and their own processes, and it is all hugely complicated” (A/NZM5). Another manager considered:

the system heart is performance management [which is] sound but can lose sight of longer-term outcomes. The theoretical heart, the heart of the system is performance, but you need to know what you’re trying to achieve and what your results and then adjust your business to as you get feedback. (A/NZM21)

Another manager outlined:

The whole theory of the whole, the way the models all supposed to hang together is pretty grand with the managing for outcomes, this idea that you set your high level goals and then you prescribe to your activities, so you have your impacts or intermediate outcomes or results that you measure those against and that should drop down then in to the inputs and that should all be reflected in the budget documentation and the estimates, should all just become this grand cascade which is beautiful and transparent and simple for everyone to hold you to account against, the kind of state their intent setting other broad outcomes, in practice, it’s a hugely complicated model to really make sense of. (A/NZM20)

These findings highlighted that in A/NZ without a country plan, there appeared to be gaps in clarity and transparency of the country’s overarching priorities. In addition, these gaps may be making it difficult for regions and communities to participate in planning and contribute to evaluating local and regional progress. One participant highlighted how “I think departments can feel under a lot of pressure to just plan on an annual cycle, and you tend to lose sight of tracking progress towards long-term outcomes” (A/NZM5). Going forward, “where we need to go with the state sector performance management is that it’s about the stability of goals and the concerted drive towards them” (A/NZM15).

However, other managers emphasised that “looking at the state sector system and effectiveness is required. Performance is very agency centric in A/NZ and [there is] limited space for sector-based outcomes” (A/NZM5). I noted in my analysis that in A/NZ,

there was a limited longer-term strategic view or strategy led by A/NZ government. Reflection by some senior and more experienced managers and evaluators suggested that 'rethinking' or reconfiguring A/NZ approaches could be required. Central to this dialogue were considerations raised by one central agency manager:

Agencies are there to do what the government tells them to do. There is a difference between the agencies having a strategy and there being government strategy. How does that fit together? How does that work? And the government may not want a long-term strategy. (A/NZM18)

Furthermore, for the A/NZ Government to honour the Treaty of Waitangi principles, my analysis highlighted that A/NZ government might need to consider structural and methodological changes. These changes could include a country development approach underpinned by concepts such as wellbeing and sustainable development, adaptive management and with more emphasis on transformative results. This could be supported by an enhanced role of M&E (including a mandate for evaluation that covers strategic and activity levels), and an improved organisational culture for sector and performance and learning, and coproduction.

Increased strategic leadership, central agency guidance and country development planning underpinned by values (such as inclusion, and partnerships) may support more participatory and systematic approaches. This research highlighted that line agency managers and performance/evaluation specialists together with central agency and development partners consider that increased outcome-focused leadership and direction at central agency level might be required. This could involve expanded agency and community inclusion, and governance oversight to ensure development effectiveness and governance was inclusive, sustainable and democratic. These findings appeared to apply to both PNG and A/NZ, albeit in a different way and would require explicit consideration. More recent literature on 'new' national planning by Chimhowu et al. (2019) concurs with these findings.

### **7.3.2. Integrating and reconfiguring government roles and functions**

Two areas highlighted by research participants for potential integration and reconfiguration of roles and functions within the A/NZ government that may enhance development effectiveness and governance, which included:

1. Alignment of investment, resourcing, non-financial information.
2. Reconfiguring and embedded theory-based evaluative approaches with community involvement and feedback loops.

### ***Alignment of investment, resourcing, non-financial information***

My A/NZ research participants said the move to outcomes was led by the State Service Commission where:

The Public Finance Act, Section 40, basically sets out the requirements around what sort of information you have to report. As part of that, Section 40 talks about outcomes as being a means of delivering social policy, came in through the *Path Finder Project* which was really done in SSC about 10, 12, even probably longer, 13 years ago, and that basically looked at the movement away from ministers contracting with agencies to deliver specific outputs, as outputs are only part of the story. (A/NZM3)

However, A/NZ research participants considered that “fragmentation occurred under the system of ‘vote’ funding” (A/NZM21) as:

the Government’s books are broken up into votes, I mean the estimates are by vote, ... [for example] the Department of Internal Affairs has eight votes, there are multiple votes in the justice sector, yet there’s nothing in the model that stops you having single strategy work towards by multiple agencies and funded from multiple votes. But the fact is that you’ve got that fragmentation does make it more difficult. (A/NZM21)

A disjunction has appeared in the A/NZ state sector as funding for votes in A/NZ agencies was for nominated outputs under output classes. Yet, with “implementing for outcomes in early 2000s, there was a need to look at longer-term goals and funding and more coordination” (A/NZM16). In addition, issues on accountability for both short and longer-term outcomes were noted with “these big social outcomes that we have, but it’s so easy to, just to fill those up with fluff and then be very difficult to hold the agency to account” (A/NZM16). However, it appeared that there is pressure to focus on shorter timeframes.

Another issue highlighted by A/NZ research participants was the impact of restructuring as it:

takes away from explicit priorities and accountability. The health sector’s a fantastic example. There was a whole lot of restructuring, in fact the sector suffers from constant structural change, partly because the difficult issue underlying the health sector is that for the foreseeable future, it will consume more resources than we can put into it. (A/NZ10)

The participants suggested that greater transparency on results was wanted. However:



I think a lot of it is actually wrestling with quite what outcomes mean, I think people have got what you might call managing for results or outputs or whatever. I think there are still some people who have difficulty in what is it that outcomes means? I think we probably codified these sorts of things a lot earlier than many other countries did, with the SOI process where departments and entities are increasingly expected to focus on what are they actually achieving for the expenditure of public funds. (A/NZM16)

The research also highlighted questions over resource allocation and community participation “how are these changes negotiated? And, how are these choices and changes negotiated? And another really important point is how the community that engages with citizen consultation?” (A/NZM6)

How are those communities interacting with the financial community, the monitors, the evaluators, and finally the people who are engaging with and being the front face of the civic choice stuff? So, there’s a big connection between a whole bunch of professionals and really important roles there that are not quite integrating themselves. (A/NZM5)

Questions were raised over:

what is valued? Who is included in that? There’s a complex of reasons why in any kind of instrumental view of public policy, you can’t hold people to account for achieving what we commonly regard as outcomes – improvements in the quality of individual and social life in various forms. But the counter-argument to that which is advanced particularly by Bill [Ryan] and others, ... is that basically, outcomes are a way of directing your efforts towards a desirable goal – they’re a constantly renewing the test of what you’re doing basically and its value. (A/NZM16)

A solution put forward by one A/NZ manager promoted collaborating and inclusion of people being involved in an “ongoing conversation” (A/NZM16) by “asking stakeholders what is important to measure” (A/NZM8).

This is something which works best if all the people who are collaborating on the achievement of those outcomes, including the people outside government, the other major stakeholders – can agree on and can somehow agree to direct their efforts towards. It’s kind of this constant renewal of effort or purpose. (A/NZM16)

The research indicated reporting outcomes within three-year government terms caused challenges “as you’re not going to report on this outcome until five years hence because that’s how long it will take, [but] you have to report on it every year. So, what

you end up reporting on is: we've done this, this and this" (A/NZM9) as "the current suite of public sector documents is not smart. It's trying to do too many things and not quite doing any of them well enough" (A/NZM5).

A participant outlined "there are some agencies in the public sector who do management and outcome reporting incredibly well, the former MAF, and MSD a bit in terms of framework, not so much in terms of data. I think IRD's frameworks are very good they're getting there on data, it's going to be a bit dry, but they do that" (A/NZM5). However, a view that accountability was for the "public" and "not just ministers" (A/NZM20) was emphasised by research participants. Another manager added:

I think there's a better understanding around accountability for the public as opposed to talking about issues of your minister and not wanting to get your minister in trouble. I think getting there in terms of we're not just accountable to ministers. I think the discipline is going in a good direction. (A/NZM20)

However, participants emphasised that further attention was still required "as most agencies still need developing" (A/NZM1). The costs for non-financial performance information were also raised as "it's just quite an expensive process actually doing outcomes and performance measurement" (A/NZM4), but you can make a lot of progress by throwing some medium-term investment at something, and actually doing "long-term outcomes" and "remain sustainable" (A/NZM6).

Overall, there appeared mixed practice across A/NZ agencies in implementing and assessing outcome performance and accountability. More systematic and consistent approaches promoted or mandated cognisant of broader inclusion and participation may be required.

### ***Reconfiguring and embedding approaches***

An example of what appeared to be an accountability gap in the A/NZ state sector was outlined by one manager:

The Government in A/NZ has a thing called the estimates and, in that arts, and performance measures. And nobody takes a blind bit of notice of those performance measures and I said to the Treasury last year when I was working on some performance measures with them on one of the votes, I said how are we going to make these sort of things work? Do you guys not want to take a note of this stuff? That's not our job. Whose job is it? Oh, the audit office keeps an eye on these things. So, whose job is it? (A/NZM12)

This issue of who is accountable in A/NZ for results was further highlighted:

Treasury doesn't feel it's accountable as an agency, Office of the Auditor-General doesn't understand anything beyond output activity measures. State Service Commission is living in a [different]...world. In the [United States Washington] DC, they have the GAO there that would ask tough questions about to that extent they were genuinely subjective and being accountable for non-performance. (A/NZE2)

This issue of accountability was also highlighted in PNG at national and provincial levels, with an increased centralised role for evaluation (PNGM5) being promoted.

In PNG, we need to evaluate what we're doing and see whether we are achieving some targets, whether there are some impacts coming out, whether there is an outcome especially on monitoring and evaluation. We need to ensure that provinces can report to the government, the national government or even to the provincial administration or provincial government on what are the impacts they are achieving, for them to see whether they are really performing or not. There needs to be something to evaluate their performance to know whether you are. So, I think those are some of the things, areas that we need to strengthen. (PNGM5)

Clarity of and focusing on underpinning concepts for outcomes such as "social wellbeing" (A/NZM15) were also highlighted:

It's a challenging thing for any organisation, but when it comes to outcomes, it's trying to achieve. So, it's a tricky question. Say, what is it about – improve our value or whatever? Is what the Ministry of Health trying to achieve – better health. So that's subjective, and then what does it mean for the outcomes that we're trying to achieve in relation to that. (A/NZM15)

A concept of "incrementality" was raised as an important aspect of making change. Yet in A/NZ, this concept of "incremental demonstration of effects" (A/NZM21) raised considerations of how to demonstrate these effects' as "we have no rules on evaluation in New Zealand" (A/NZE2). Research participants noted that "other jurisdictions such as Canada and the United States" (A/NZE3) were passing laws for evaluation where solutions included embedding evaluation into strategies, and central agencies taking a more significant role in increasing accountability such as:

Obama has passed a new law for evaluation. There can be incentives in place which require agencies to have these sorts of things. And then it holds people to account. (A/NZM4)

### ***Is changing the approach and rules for evaluation a solution?***

A solution put forward by one A/NZ central agency manager was potentially “changing the rules” for examining the effectiveness of the government (A/NZM21). This could include being more prescriptive as has been done in other areas such as “how capabilities systems work” (A/NZM21). Rules promoted for change included:

relating to the appropriation of resources is where this comes through. The government has not prescribed how policy is written; it could do of course. But it does prescribe how capabilities systems work. So, most of those rules are written by accountants, which have got all the limitations. Some of those rules, I think need to be developed further by people who are more interested in national impacts in our society. I think that would help a lot. (A/NZM21)

Using approaches “creatively and dynamically” may assist with “a more embedded form in our society” (A/NZM21). A supporting example was provided:

Using a diabetes example, having gotten that on to the agenda what are we then really going to about it and how seriously are we going to be able to deal with Type 2 diabetes, what are we really going to do about it? And can we, have we got to approach these things creatively and dynamically? Obviously, in societies when there’s times of crisis, people do look to countries like Israel and Finland which are countries that respond to crisis. But we need to find a way to deal with this more, in a more embedded form in our society. (A/NZM21)

The research pointed to an integration of functions that may be required “as I’m kind of concerned about the lack of integration of the different functions. That they should be complimentary. They should work in a tight understanding of how this organisation is here to achieve these sorts of results” (A/NZM21). However, “I get perplexed about why there isn’t that line of sight in the information which then allows you to use evaluation intelligently – I’m not sure we use it intelligently at the moment” (A/NZM21).

Moreover, an A/NZ manager outlined that “approaches for evaluation need to vary. Evaluation to be effective has to be as close as possible in time and space to what is being evaluated” (A/NZE8). The more “formal” type of evaluation had timeliness concerns expressed, including:

There’s this kind of formal scientific evaluation which has the strength at least that you have a clear hypothesis that you’re trying to test, to what extent has this desired state been achieved, but nevertheless occurs quite often. It tends to ask questions that you can answer rather than questions that you should answer. It tends to be remote in time and space from the actual events that you’re trying to evaluate so, you get a very professionally done evaluation, it comes out two years

after the event and what happens? The people you are reporting to say, 'Oh no, we've moved on since then,' it's all completely irrelevant. (A/NZE2)

A form of "collaborative evaluation" (A/NZE2) was promoted. A solution outlined by one A/NZ evaluator included:

You do have to be able to create an environment where people see better information about what they're doing not as something that poses a threat to them but provides them with an opportunity to do better basically, and where they don't feel that somehow they're being chased if things don't work out the way they should. (A/NZE2)

The focus of evaluations was also questioned where "often when agencies do evaluations, they'll do evaluations of a provider for example. When they're looking at intervention, they'll look at the providers that they're funded to do an intervention, but they typically don't look at themselves, or quite often they don't" (A/NZE2). A concept of "joint accountabilities" (A/NZE6) was promoted:

it's hard for individual ministries with their own ministers to work together because they've got their own accountabilities to the ministers, but there's a bit of a movement towards joint accountabilities. ...Internal entities are working towards common outcomes, I think there's people that are really keen to do that, and there's quite a few frameworks currently being developed within various business units that try to focus on what those business units do in relation to their also try and work with other business units. (A/NZE6)

### ***Role of central agencies***

Research participants consider that central agencies in A/NZ need to define the role and function for evaluation as:

Public sector's kind of centrally driven so you have the key agencies of DPMC [Department of Prime Minister and Cabinet], Auditor General, State Services Commission, and NZ Treasury. It's those agencies which need to sort of really start to drive evaluation into it and also the government of the day needs to be able to see how, in fact, it's going to help them politically as well. (A/NZM8)

This research interviews showed there were initiatives underway focused on "performance improvement in public sector using a state sector model" (A/NZM18) being undertaken by SSC. The Performance Improvement Framework (PIF) scope included "for the core business areas, we're looking at effectiveness and efficiency and you demonstrate you're doing this effectively and there are various areas that you look at and efficiency as well" (A/NZM18). For the central agency performance initiatives:

the driver was looking at how the central agencies work together and saying that we don't have a combined view in terms of agencies. For example, agencies would talk to Treasury about something, and they would say this and we would, then they'd talk to the SSC always say something completely different, so it's saying ok, that's confusing for agencies. So one of the performance improvement framework objectives was to create a common language for performance, so have we achieved that? Part of that coming towards the common view of performance with the central agencies so the people from SSC, the Treasury and DPMC are in the same room talking about an agency using a common framework hearing the findings of the review team, it's all coming together as at least a more unified view of an agency's performance. (A/NZM17)

However, a central manager emphasised that "one of the things that we find is you can often see really good, documented policies and processes, but at the end of the day, it's the application" (A/NZM18).

They try, the Treasury and SSC and put out some high-level guidance, however for, departments are so different it's hard for them to give out kind of to all departments, so it's probably of limited use. I think it's the Office of the Auditor General has been a little bit better, sort of seem to be leading this process. I don't think they wanted to lead it, but I think that they ended up leading it. They have put out some better practice guidance so they've actually put examples of better practices out there so the sector can look at them and see what good practice might be in their opinion and leverage off that. (A/NZM16)

### ***So, who is accountable for outcomes?***

It appeared from this research that some efforts were being made by A/NZ central agencies to improve accountability guidance and documentation.

I think the agencies are doing a lot of outcomes work and developing a lot of outcomes thinking and especially off the back of the State Service Commission kind of direct for agencies to be thinking about outcomes rather than outputs, but there's a real big grey and fuzzy area, I think around correlation and what that agency's accountability is and that sort of thing. (A/NZM16)

Yet, A/NZ line agency research participants advocated for central agencies taking a greater role as:

...arguably for agencies like Audit NZ perhaps and NZ Treasury, if monies dished out then, you might be looking at how is money used and prove what contribution you've made. In a sense [like] what's done with the *Whānau Ora* outcomes

framework, which [includes agency level outcomes] and [that] doesn't actually seem to appear in that many frameworks (A/NZE10).

However, participants considered there appears to be an accountability gap for outcomes in A/NZ and suggested that two limitations may exist including: (1) level of institution accountability for outcomes; and (2) central agency reporting and accountability requirements.

A key finding from my analysis was that there is not an explicit role for evaluation in A/NZ as the focus in central agencies is on review of the chief executive as outlined in the PIF documented approach.

Under the actual sort of heading of managing for outcomes in the SSC than there perhaps was in recent years, up to comparatively recent times a lot of the commission's focus really has been on measuring the performance of the chief executive almost as the proxy for the performance of their department. (A/NZM20)

Furthermore, there was a potential gap over who is accountable for the extent of results in A/NZ. In addition, this appeared to raise questions on the appropriateness of devolution in A/NZ as:

the chief executives in A/NZ have compared to most other countries, substantial autonomy in terms of responsibility for production of whatever you want to call it, results, outcomes, against their budget, responses so they have a sort of slightly awkward triangular relationship with their responsible minister or ministers to whom they're responsible. But also, a third leg of that is their responsibility to the commissioner as their actual employer, and then chief executives, in turn, are employers of all their staff within their organisation. (A/NZM16)

This pointed to a potential direction and accountability issue for development results:

What drives performance in A/NZ? Unless you've actually put it at the heart of the thing, no one's at a seminar or workshop where Treasury are looking at the next steps of a public management system. And I said, well you've got to make your mind up, is it inputs? Is it outputs which is AG4 or is it outcomes which is managing from outcomes, what's the premium mobile, what drives the system, what drives the expectations on people, what drives performance assessment? (A/NZM15)

In turn, this led on to me reflecting about A/NZ context – what is the way forward to get results? Changing the focus? One manager reflected:

So, what's the government role in achieving that and we sort of look at outcomes in terms of capabilities, which I think Treasury is moving towards. And I think that approach, like we've used are a promising approach and measuring any outcomes. I think it doesn't mean we don't have to, but it means that we focus on what difference can the government make in terms of wellbeing and health. It's about democratic rights and freedoms, and it's about people who value freedom and the right to health of education, economic, stuff like that and being safe. (A/NZM12)

Changing the focus to more on results may require explicit consideration:

It's a tricky space to be in because I don't think we traditionally, we thought about outcomes in that way. People notice in public health, and we try and look at what we're trying to achieve. It's a useful one for Treasury too because it's come up because of the point of having is not money, it's about the benefit of all A/NZ citizens, so in terms of wellbeing and outcomes and all that. (A/NZM16)

In 2016, PNG passed legislation mandating a central agency role and accountability function for evaluation within DNPM. A country approach for evaluation was being developed aligned to the MTDP to support monitoring, evaluation, analysis and reporting of progress within the sectors and at a country level. Schwandt et al. (2015) advocate that evaluation can play a role nationally in assessing whether progress is equitable, relevant and sustainable. PNG/s centralised mandate for evaluation may enable these analyses and adaptive responses to be undertaken.

Findings from this research demonstrate that further attention and explicit consideration on underpinning development and evaluation concepts and goals may be required by central agencies, line agencies, regions, communities.

### **7.3.3. Increasing focus on transformative development results**

The findings highlighted that the "focus needs to be on outcomes and longer-term goals. It's not only about having clear outcomes and having expenditure and various outputs and impact behind that, it's about achieving long-term social goals as well, social or economic goals whatever they happen to be" (A/NZM4). Working in a partnership was promoted, and "building the capability, so people are comfortable to do that [work together] and recognising where they need to kind of slightly also push and help everybody involved to move the next step" (PNGDP).

Communicating about results was emphasised involving "succinctly being able to articulate what we're actually achieving. It's a big challenge though, and it's not going to be solved overnight" (PNGDP). Using frameworks to link inputs is highlighted:



When we're developing our frameworks, we're thinking about getting beneficiaries' input into assessing progress. So, it's there, but I think we could be doing a lot more. We're probably taking smaller steps at the moment, just trying to get the basics in place before we start running. (PNGDP)

A PNG participant highlighted that "results, accountability, transparency have all popped up as key principles.... And also, there's an action plan that's come out of Busan [High-level Forum] around statistics as most people recognise that statistics and data is key to actually producing evidence around what we're actually achieving out there" (PNGDP). Challenges with aligning inputs multiple partners were also noted by development funders "as you've got sector plans and how you are aligning with the country and other donors" (PNGDP), which can require negotiations. In addition, contextual complexities for funding flows were highlighted in PNG, particularly for sub-national systems.

I think the big thing for PNG in development is districts towards and the funding flows, the focus now is really on services. You've got 89 districts, 306 LLGs [local-level governments] and 6000 plus wards out to villages. To get the services out in the different sectors, the capacity, for example, health plans – what is the funding mechanism? Is it going direct out to districts? LLG is the funding within that. (PNGM6)

My research found that the multiple layers, system components and stakeholders at different levels – national, provincial and local – all add to the complexity for management, development results and governance.

### ***Importance of context, relationships and outcomes***

Several participants highlighted the importance of understanding country contexts, constitutions and relationships in both PNG and A/NZ. In terms of the latter, flexibility in relationships were emphasised:

Managing relationships and many of their expectations of both what we're accountable to our stakeholders and to our taxpayers etc. I think there's also different ways of doing that, of telling our story and how it's not necessarily all statistical. We do have statutory requirements that we have to meet but it's around about having that good relationship with Audit NZ, Treasury. Its awareness and understanding of where we're coming from, some of the challenges, where we're trying to go with it and them understanding. So, what they ask is also realistic and achievable. (A/NZM8)

In A/NZ, understanding how outcomes related to staff in different roles and whānau [families] was particularly emphasised. Working towards a convergent view is advocated with “thinking about outcomes and trying to unpack the impact of government and peoples’ experience of it – you have to start with the base which is trying to do outcomes. Focus on what difference can the government make in terms of wellbeing” (A/NZM14). Underpinning concepts were promoted such as wellbeing and health.

It’s about democratic rights and freedoms and it’s about people who value freedom and the right to education, economic and being safe. So it’s a tricky space to be in because I don’t think we traditionally thought about outcomes in that way. People notice in public health, and look at what we’re trying to achieve. It’s a useful one for Treasury too because it’s come up because of the point of having is not money, it’s about the benefit of all A/NZ citizens, so in terms of wellbeing and outcomes. (A/NZM9)

In relation to development partners, “the best laid plans of a programme will founder on a poor definition of the environment that it’s going in to. You can build a dam and give them electricity but if they haven’t got jobs to buy the electricity it’s not going to lead to anything” (A/NZM16). Local leadership was emphasised to provide local context and input.

Some provinces are more forward because they have good leaders, both politically and socially. If you don’t have a good leader, if you don’t have a good political will there then you have a problem. So, in some provinces we have a problem, there is no leadership in the province. So, leadership is critical and one of the things that we are working on now is we want to make sure that we work on mentoring and coaching of our provincial administrators...to be able to lead the province and then coordinate. (PNGM10)

Leadership was viewed as “critical in every organisation and I think we’ll begin to realise that we need to make some changes” (PNGM6).

### ***Monitoring, evaluation and data***

A key part of assessing development progress requires active monitoring and evaluation, analysis and using evidence in decision-making. One participant stated “there is still some thinking to do about the whole systems on this. The M&E role within this needs to be thought about in terms of from a national planning perspective looking at a project level, outputs, outcomes, activities” (PNGM5).

They [funders] really want to see progressive checking. There’s some funding going straight to districts on, and the trouble is that, or donors if you fund everything upfront there’s no incentive for actually funding. So, to make sure

things happen, you can look at progressively funding which donors would come under. And then that gives more control about controlling based on performance, on activities, outputs, outcomes and impacts. (PNGM6)

Both PNG and A/NZ research participants advocated “having beneficiaries involved in planning, monitoring, evaluation and decision-making, and increased controls for performance and accountability” (A/NZE6). This can include

quantitative and qualitative [data], want to know how it is for the beneficiaries – At the moment Planning’s still not ready, currently restructuring how to do this. It’s a priority, the national plan. The work plan activities report back to us, so they see the funding, the whole system coming through national planning as the coordinating agency, but the M&E division got disestablished, but they are now restructuring to look at this, so a function to measure progress. Central procedures, not just policies but actual planning and the role, so it’s moving back in from a policy but also performance controlling because, under the direct funding to districts, there’s been less control which has been a concern about holding people accountable for the performance. (PNGM10)

Likewise, participants emphasised that access to data and systems were wanted in the A/NZ state sector context. Hence, for example

We capture a huge amount of information but being able to articulate, simply some key results that are delivered. I think we’re still trying to grapple with. Given the current climate, there’s a lot of focus on cost-effectiveness, efficiency and one of the easier ways to show what we’re doing is how many outputs are being produced? That’s not to say that we shouldn’t also focus on the outcomes that we’re achieving as a result of inputting and outputs. (A/NZE9)

Having increased participation in planning and M&E was also wanted by stakeholders with coproduced outcome-based data. These approaches will require further attention by central agencies.

#### **7.3.4. Summary**

This research highlighted that participatory country development results planning, portfolio funding and integrated role of M&E could become country and sector level requirements for results and accountability at both strategic and activity levels. The participants wanted more inclusion of key groups in development planning and governance, alignment of investment, and embedding adaptive evaluative approaches to occur. This will require increased central agency leadership and guidance to support alignment across the sector and line agencies, private sector, NGOs and development partners.

## 7.4 Emergent concepts – requiring inclusion

The following section outlines the emergent research findings requiring inclusion associated with the *wider infrastructure* layer (Table 13). The findings were narrated under the emergent concepts.

**Table 13. Emergent conceptual areas requiring inclusion: wider infrastructure layer.**

Concepts		Evidence
1	Increased governance and systematic community inclusion are required.	The qualitative data, RT and AR displayed by A/NZ line agencies ranged from fair, limited to no systematic or explicitness inclusion of key population subgroups which are required to enable more relational and inclusive development approaches.
2	Values-based country development results planning, portfolio funding and the role of M&E are country and sector level requirements for results and accountability at both strategic and activity levels.	Evidenced by QF. Having guiding values can support more inclusion of key groups in development planning and governance, alignment of investment, and embedding adaptive evaluative approaches with centralised central across sector and line agencies, private sector, NGOs and development partners.

### 7.4.1. Increased governance and systematic community inclusion required

The accountability document review rating *for inclusion of key subgroups* (ethnicity) Criterion 6, Table 1) ranged from fair to no systematic inclusion. Yet, inclusion of Māori in planning and development is a country principle included in the Treaty of Waitangi. My analysis showed that more governance and systematic community inclusion may be required in A/NZ to enable more relational and inclusive development approaches.

One A/NZ manager outlined how challenges in prioritisation occurs:

Government to ministerial priorities, it's one of the harder things is that you get a bit of a tension...there's things the Ministry wants to achieve, and we've got outcomes and things we're aiming to achieve. And ministers have different things, but they decide that's where funding's going to go and we're going to do this first. I guess helping to sort of even be able to push back on ministers and go if you do that then we might not get there, or we might not be able to do this. And it helps a lot to hear those conversations. (A/NZM18)

Conversely, the PNG planning system demonstrated more systematic regional inclusion reflecting the mandated country planning framework. These findings appeared to have significant implications for development planning, ownership and sustainability of development results and data, ownership, and ultimately citizens' democratic rights.

Whereas, in A/NZ, there appeared to be confusion on who defines the outcomes for communities and local level, and who was included in evaluating progress as:

we have this interesting sort of scenario in A/NZ where government then decides that there are big outcomes that it wants, the government of the day wants these big outcomes to happen. And then through the machinery of government which the bureaucracy we have all these programmes and policies, and which are pots of money essentially. (A/NZ11)

However, a mismatch was noted:

There seems to be a mismatch between community sector evaluation and government evaluation. We need to get that a lot of the community sector-based evaluators have got a lot of value as they've got a lot that they can teach us, [that is] us being in our perspective right now of being based in government. I think there's a lot we can learn from them. (A/NZE11)

An A/NZ evaluator considered how to bridge these two paradigms:

It's actually really important and I think there's a bit of a gap at the moment. Government evaluation people coming through have that insight that they actually even went and spent a week sitting alongside a community provider maybe they would see it's quite a different world as we need to remind ourselves that those providers don't exist for government, they're there because they're passionate. There are two paradigms happening – about what they do, and then they see a reason for it. (A/NZE11)

This disjunction between government priorities and people on the ground was also noted in PNG in relation to education:

The teacher's point of view that they did not agree, but they just went ahead. And then, the teachers were overloaded, the teachers who were at the start of the outcome basis had no time to sit with kids when the classroom is filled with students. (PNGM7)

This finding raised questions on who defines and confirms the outcomes and planned activities. A participant emphasised the value of "working together from a shared understanding" (A/NZE6) as:

you need to think about what people understand as an outcome. If somebody on the ground the beneficiary, the recipient of the service, what they think is an outcome and what a public service thinks is an outcome or the delivery person thinks is an outcome if they don't have that same conceptual basis then you're

not really going to have people being able to co-produce outcomes for themselves. (A/NZE6)

#### **7.4.2. Country-led development and aid effectiveness**

A PNG participant highlighted how country coordination and an inclusive process was promoted for setting development goals at the broader contextual level:

It comes down to PNG having a PNG coordinating committee from the Paris Declaration, saying here are the plans for M&E. Here's going to be an internal reflection meeting later in the year where basically you can jointly track progress against the targets, against the indicators, against the plans. The role of the database, the business plans, how they can be customised, so this is where the focus is for PNG this year. (PNGM5)

Participants also saw that planning undertaken collaboratively by development partners was derived by:

....monitoring of projects, national planning together. The other thing is you've got the organic laws, the law that it really does, policies progressing development partners, roles of NGOs, private sector how that all works together. And looking at say whether the health [sector] is working in the how does that roll out in terms of health committees right down into all provinces right out to the people? (PNGM5)

In addition, extending this collaboration was also advocated for evaluation as:

Donors are doing their own evaluations, but we've got to be able to get the information that comes in. So, this is where the database about missions coming into the country, how that's going to benefit and how is it all linked together? So, PNG can prepare, and it can be used systematically. (PNGM5)

PNG research participants emphasised using systems approaches for "aligning plans and programmes, agencies, provinces, processes and budgetary alignment. Financial management needs to be strengthened which they're underway on, and integrated information, which is led by GoPNG, so it can be used by nationals but also donors" (PNGDP).

The goal is to have that financial system really strong – systems and processes. And then to align development partners to national plans, organic law and all working towards particularly for example, in the health sector, towards the health indicators. (PNGM6)

When considering how NZ Governments' aid programme interfaced with PNG systems, Wood et al. (2008) observed that: "there is scope to enhance NZAID's

implementation of the Paris Declaration further, particularly alignment with country systems and processes” (p. 108). My analysis showed by using systems approaches, collaboration and coordination can be improved nationally and with development partners which appears to enhance development effectiveness.

#### **7.4.3. Country development results planning, portfolio funding and role of M&E**

My findings noted that clarifying approaches for country development, results and role of M&E was underway with PNG and by some development partners as “at the country level there’s quite clear requirements around results, accountability. At the activity level, the expectations are there, but it’s less prescriptive about having to meet at least once a year, but there are regular conjoint discussions around results” (PNGDP).

PNG research participants outlined that a rationale for aligning development partners was that “different requirements from different donors and different government agencies can sometimes paralyse the countries because they are different. So, it’s a roundabout kind of moving towards closer alignment” (PNGDP). This process included:

early discussions around what they’re using and aligning it with the other donors that participate in the Pacific. So, we’ve had a couple of early good conversations with them around the need to do that, so it doesn’t become a burden on the countries. With having to start and have early discussions around building not just the monitoring and statistical but the evaluation capability across the Pacific. That’s not going to happen this year but it’s around about starting that conversation and getting the commitment and developing. (PNGDP)

However, a recognised challenge for a development partner was:

being able to articulate the results of our work. To build a system to enable us to do that. There’s a bit of a trend out there at the moment around developing some standard indicators which they apply across their aid programme which enables them at the end of each year to say ... although it’s only a narrow focus and it tends to be sort of output focused, it does give a flavour to the public of what is actually being delivered. (PNGDP)

International efforts were noted “as basically there’s been some effort worldwide through DAC and other places to try and make them [funders] good citizens in the international donor community” (A/NZM16). Considerations outlined for evaluation included:

Evaluation depends on the quality of the management process you’ve got in place. The kind of higher-level evaluation has got to be important, and what we

tried to do was to think of the whole thing in terms of those four capacities to state that Marylee Grindle had identified. Essentially, evaluation is underpinned by all four of them. (A/NZM16)

Grindle's (2004) four capacities for a state included:

You have to have a state which has the capacity to implement to start off with, which means having basically effective control of its territory. You have to have a state which can understand itself, it's got to have the cognitive capacity to understand what effect's it's having which is a whole system of communication with the outside world and within itself. You've got to have a state which has the technical capacity to implement as well as the power to do so, which is again the quality of your management... And you've got to have a state which has the political capacity, basically the incentive to respond to the signals that it's getting to do better. (A/NZM16)

Having "a functioning state with these capacities and governance" was emphasised:

as if you don't have each one of those conditions in place you cannot have effective evaluation really because evaluation depends upon an effective functioning state which is able to sense what it's doing and respond effectively to it and has the will to do that. So that's where the governance comes in. (A/NZM16)

Moreover, a role for evaluation within governance was also promoted.

There's a tendency in evaluations to kind of float it free from all these other aspects of governance, to see it as some kind of God-like function which will provide a summit of judgement on the performance of the state in some area, which is useless unless in fact the people who make the decisions and implement them really want to know what's going on and to do something about it. (A/NZM14)

In addition, a role for "effective participative evaluation" (A/NZM16) in government was promoted where inclusion and participation were actively undertaken by regions, communities and providers. An example was given by an A/NZ manager:

The thing I discovered about Samoa is that nobody is reluctant to complain. We visited a lot of villages and we got some very straight and direct feedback about the quality of service they were getting from the three key departments basically – health, education and agriculture – worried that there was this kind of growing distance between the government in Apia and the villages, said that basically, Samoans view the government in Apia as their servants, the people that are



responsible for making sure that they get the service they deserve or if they don't get the services they deserve, they'll ...go to Apia and complain about it basically. And that's something I think with the relationship between the villages and the Western-style government which dislikes that sort of thing. (A/NZM16)

#### **7.4.4. Governance, management, accountability and democracy**

My analysis displayed different perspectives on the role of the state in A/NZ, which were highlighted in the findings. One manager outlined:

the government is saying we want to see improved systems performance. Say here are some targets and goals and we'll measure ourselves against them, and it makes sense to have the same measures for external accountability purposes as for internal purposes, whether it's about scorecard or dashboard, but whatever the leadership team or senior management uses and also what teams use and talk and measure their performance against. (A/NZM16)

Another A/NZ manager outlined an issue over the role of the state:

There's a kind of inherent tension between letting politics, and the democratic process do its thing, but there are businesses delivering serious services that need to be sustainably maintained and can't just be allowed to fall over, and essentially [that is] why constitutionalists and managerialists clash. (A/NZM21)

Moreover, the need for an expanded role of governance was emphasised "as there does seem to be ... an expression of the total governance and the flow of information to Parliament ...[which] means that we often avoid the unspeakable or the things that are most the part of the public debate" (A/NZM15). This led to considerations on the role of choice over the use of financial resources for outcomes and transparency.

There's going to be a lot of financial challenges and so careful choices and better knowledge about how we understand our services and what they're achieving and what we can do with them I think is going to start being fundamental. So even if the system and its formal mechanics doesn't prescribe it at the moment, everybody's going to be looking there. Circumstances are going to force us, I suspect. (A/NZM15)

An approach of looking at "themes from a government perspective in A/NZ" (A/NZM12) was promoted as:

it's challenging. I think there's a lot of interest in outcomes and government performance and I think the main thrust of it has not been looking at the dual agency or initiative outcomes but what they mean in total from a government perspective. So, what are the current themes? Not having outcomes in

educational health as what do they mean in terms of social welfare and whatever? With the social sector, it's on other parts of the Ministry, what we focus on now is the work of the Ministry. It's about what does the work of justice, health, MSD achieve in total rather than this is education, that's health, and how does work? (A/NZM12)

One PNG participant outlined that one development partner agency had “revised business rules to clarify approaches from strategic down to activity level” (PNGDP). My research highlighted another potential focus area for standardising approaches:

There's going to be a greater degree of standardisation which becomes a mandatory step. At the country programme level, we concluded a joint commitment for the development at the end of last year in October. And each of those needed to have results framework ... These initiatives are being undertaken by countries and development partners and are aimed at enhancing development outcomes where it is very much about looking at the development outcomes that are, the critical ones... as a whole. (PNGDP).

However, in contrast to mandated approaches, several PNG participants raised issues of mixing political appointments in the management of agencies was noted by one PNG manager. “Our government system I feel it's a little too flexible. I believe the public service should be kept away from politics. Most head of departments are politically appointed which is okay, but that should not have any impact on what we do – is important” (PNGM6).

In addition, my analysis showed that a country plan developed with inclusion of sectors, agencies and development partners might enhance alignment, country priorities and potentially increased country ownership.

Having country plan MTDP, you'll hear from a number of donors is that while it's a very important document and there is actually a clear commitment to it pretty much all government entities align their planning with. Even in, on Bougainville which is autonomous still has to. The difficulty that a lot of donors have is that it was actually developed without consultation with donors and perhaps more significantly there were a number of line ministries that weren't sufficiently in our view involved in the formulation of it. (PNGDP)

In addition, orientating sectors to country plans also appeared as a key process to align inputs and contribute to enhancing development effectiveness.

We have aligned with the government because we're working on the basis of this sectoral strategy and the sectoral plan. And there's a degree of reorientation now

involved to the MTDP. So obviously we recognise just how important it is and we'll do our best to do that, but I think it is one of the realities. (PNGDP)

These findings highlighted that an increased role of central agency might be required, ensuring accountability and performance occurs with aligned financial systems and integrating M&E systems.

I think DNPM is trying now to do more, to get that alignment dialogue going again and increasingly trying to push hard for donors to engage with it and to play a coordination function. (PNGDP)

The research evidence showed that both A/NZ and PNG central agencies were either expanding or considering how and if they could increase their central agency roles to provide more country direction and focus in areas such as wellbeing in A/NZ and sustainable development and service delivery in PNG. These changes may lead to more inclusive and improved development effectiveness and governance.

These findings and further reflections led me to consider that maybe some groups of A/NZ's citizens are rather accepting of the capability and capacity of how A/NZ government's operate. Are the strong biases of the predominant European population based on colonial legacies maintaining the power and control to keep their advantages within society? Could a collaborative country plan for A/NZ really assist in building a more inclusive, equitable and sustainable way forward that includes a focus on the wellbeing of its citizens as promoted by Dalziel and Saunders (2014)? I observed that the PNG country plan had assisted in dialogue and transparency of development goals. Yet, for countries to really progress against development goals more equitably, strong representative leadership, and increased governance and accountability oversight may be required to ensure public funds are used effectively.

## **7.5 Chapter summary**

The research question driving this research is: 'What evidence from within *ideas, interventions and frameworks; institution; and wider infrastructure* settings in A/NZ and PNG demonstrate what works for whom in relation to results frameworks' and associated management systems? This chapter focused on the wider infrastructure context.

This research highlighted that a coordinated country and sector approach and systematic regional inclusion in PNG, and some A/NZ sectors may contribute to improved development effectiveness with goal prioritisation, and enhanced services and environmental awareness. Increased strategic outcomes leadership, integration of development, evaluation and management through use of a state sector systems business model, a more explicit function for evaluation, an enhanced central agency role

with increased use of results and outcomes frameworks, and alignment of development partners (national, NGO, private sector, donors) to country priorities and increased accountability were all considered as important aspects by research participants.

The research showed that line agency managers and performance/evaluation specialists together with central agency and development partners supported and considered an increased outcome-focused leadership role at central agency level was required with agency and community inclusion and governance. A more explicit focus on development theory and approaches and an increased government leadership role with community inclusion may enable more systematic, adaptive and democratic development approaches. This shift might support an expanded implementation of services and interventions for people in the regions and communities.

However, increased attention was required in both countries to expand these approaches. PNG and A/NZ and other countries can learn from aspects of these approaches such as mandated country-wide inclusive regional approaches and an expanded role for central agencies in development planning, portfolio-based funding and learning and performance as this research showed enhances development effectiveness and governance. This research showed how vital an explicit inclusive approach was to support democratic rights for citizens in more inclusive sustainable development and governance.

These findings were significant as they challenged both A/NZ's central agency devolved managerial approach to the line agencies and PNG's funding oversight to line agencies for regional service delivery. Questions were raised on the nature, role and approach of the state, governments and public sector and development partners. If the articulated goal is inclusive sustainable development, then the current paradigm of development, management and evaluation constrains inclusion, governance, accountability, and sustainable development.

In Chapter Eight , a critical discussion was undertaken, identifying critical gaps and suggested areas where different concepts than those currently used needed to be reconsidered for development, management and evaluation knowledge fields in relation to country, sector and programme results frameworks, and associated management systems and governance. The identifiable good practice concepts were then assessed to what extent that these emergent knowledge concepts and existing concepts required inclusion.



## Part C: Discussion and conclusions

### Chapter 8: Concepts linking frameworks, management systems and principles

My thesis was that current evaluative approaches to the management and governance of development, particularly in the context of the SDGs, were insufficient to ensure robust, inclusive, and effective results. The overarching research question that I focused on to undertake this theory-building research using a new typology as the conceptual frame was: *What are 'good practice' evaluative principles and concepts that may enhance country, sector and agency development, management and governance knowledge and practice in different contexts?* Across all three of the contextual layers, certain common themes emerged from the research findings. In this chapter, I engaged with these themes to examine the nature and role of evaluation in development, management, and governance within the current paradigm. As part of my theory-building, I then suggested a possible reconfiguration of these themes within an emergent paradigm shift for management and strategic evaluation that may encourage more inclusive and sustainable development practice.

This chapter contributed to addressing four of the sub-research questions: What evidence in A/NZ and PNG demonstrates what works for whom in relation to results frameworks and associated management and governance systems? (RQ1); How and in what ways can conceptual links be identified between frameworks, management and governance systems and good practice principles? (RQ2); How and in what ways can country, sector, and agency strategic results frameworks and their associated management and governance systems underpin the development of good practice values and principles? (RQ3); and How do the emerging good practice values and principles in country, sector, agency and programme systems work to enhance development effectiveness and governance for development (national governments and partners), management, and evaluation? (RQ4)

This chapter was divided into three parts: First, a critical discussion of the findings that relate to RQ1 was provided for each of the three contextual layers in my conceptual research typology – namely (1) *ideas, interventions and frameworks (ideas)*, (2) *institution*, and (3) *wider infrastructure*. Second, identifiable good practice concepts from this research were assessed to determine the extent to which existing and emergent concepts, values and principles are relevant to development effectiveness and

governance (RQ2). Third, a proposal was put forward for reconfiguring strategic evaluation into management discourse (RQ3 &4).

## **8.1 Critical discussion**

My initial findings indicated that the dimensions included in the impact models (Figures 2 (initial) and 7 (updated with emergent conceptual typology of three contextual layers) were limited. Research participants noted additional relevant dimensions to enhance development effectiveness and governance that were not captured in the OECD (2005b) guidance. In addition, based on my scoping phase findings, the importance of interpersonal relationships and individual capacities, acknowledged by Pawson (2006) in his original model, appeared to have relevance within each contextual layer. Thus, a new research conceptual typology was developed comprising three contextual layers with interpersonal relationships and capacities included within all three levels (Figure 5).

My analysis against the impact model dimensions identified critical gaps in the ways in which development effectiveness was currently understood and suggested areas where concepts other than those currently used need to be introduced and considered and where some existing concepts which required more attention (refer Appendix G for a full summary of identified concepts from this research).

As part of my theory-building approach for this research, I suggested that embedding evaluative approaches at strategic rather than operational levels was crucial as it leads to a greater emphasis on the relationships and capabilities that underpin the strengthening of management and governance systems. Fourteen themes (see Table 14 and Appendix G) were outlined that required further attention or inclusion including six themes in the *ideas* layer, four themes in the *institution* layer and four themes in the *wider infrastructure* layer. Each theme was discussed with reference to the wider literature and supporting quotes from research participants where appropriate.

### **8.1.1. Layer 1: Ideas, interventions and frameworks**

In terms of the *ideas* layer (which encompassed interventions, associated frameworks, interpersonal relations and management and evaluative capability), three concepts that did not seem to be addressed well in existing agency practice in either country were: (1) inclusion, transparency and accountability and managerial capability in relation to the use of frameworks, (2) the explicit engagement (or not) with sustainable development, and (3) the role of evaluation.

## **Frameworks**

The first point of discussion was that an increased use of results frameworks (or outcomes frameworks) in A/NZ and an increased focus on the PNG country plan (MTDP) provided greater transparency of intended results and a basis for dialogue and reporting progress. However, two issues were reported in relation to these frameworks: first, participants noted the importance of who is included in the dialogue in developing these frameworks and, second, the question arose of what capability managers and communities need to develop and participate in the use of these frameworks.

In relation to developing these frameworks, my findings showed that participants regarded frameworks as specialist tools used by (internal or external) evaluators rather than capabilities required by agency staff. Managers in A/NZ noted their lack of capability to develop and use these frameworks as a management tool for planning and reporting. This has led to the outcome or results frameworks being developed by A/NZ evaluators or performance management specialists and then a limited ongoing use as practical management tools. There was also a perception that critical external stakeholders were excluded from participating in the ongoing use of these frameworks.

However, an emerging, positive shift evident in some agencies, such as MFAT in A/NZ and in DNPM in PNG, was for managers to be beginning to use sector and results frameworks as dialogue and reporting tools. This highlighted partnership and goal alignment linked to these strategic frameworks and external demand. PNG's MTDP was an example cited by participants where the country framework was used to align government and development partner inputs. In both settings, there were guidance and capability support provided for managers. In A/NZ, emerging use of sector frameworks (for example, in social development and justice) appeared to enable collaboration and alignment for shared results across multiple agencies. However, despite this shift at management level. There appeared to be little being done to clarify the agencies' work programme on the ground. The frameworks did not appear used down to the level of operational and community engagement. So, while results frameworks may be being used more, the question of inclusion was not resolved.

A further finding of this research was that key stakeholder groups (including Māori and Pacific peoples in A/NZ, and line agencies in PNG) were not effectively included or invited to participate in planning decision-making. In the A/NZ setting, this breached Treaty of Waitangi principles and raised democratic considerations about equitable access to services. In the A/NZ system of devolved managerial responsibilities for government agencies, an underpinning assumption was that inclusion occurs because the Treaty is a constitutional agreement recognised under international law. However,



my research highlighted that potential accountability gaps occur between central and line agencies and the assumption of inclusion often does not hold. In addition, assessing whether targeting of services and interventions was equitable depends on institution capacity to access and use appropriately segmented data, but this appeared unsystematic. My qualitative findings suggested that values of inclusion, partnership and participation may require an explicit mandate from A/NZ central agencies to ensure more systematic inclusive and transparent approaches are undertaken. One A/NZ manager noted: “the good thing about having a framework and trying to use a framework is that you actually make it explicit about how you’re valuing things” (A/NZM9).

Whereas, in PNG, the mandated inclusion of provinces under Organic Law in planning meant that provincial priorities were more systematically considered as provinces have some input into budget allocation and decision-making. However, my research highlighted that the systematic inclusion of provincial needs and coordination across sectors did not occur in A/NZ. Participants considered that non-inclusion and participation have contributed to inequalities for Māori and Pacific in areas such as health and education with disparities for accessing services in regional areas. These issues remained, and my findings pointed to a need for central agency intervention in A/NZ to ensure more equitable and inclusive processes occur into planning and decision-making.

In PNG, increased central agency-level efforts to integrate evaluative approaches into MTDP planning and budget processes were underway to enhance measuring, reporting and adaptive budget allocation. It appeared that funders and donors in PNG were identified at a central planning level. This pointed to GoPNG, using country and sector approaches, to effectively reduce implementation fragmentation across its multiple partners. Use of a range of theories of change has helped align multiple sector priorities with development goals. Participants outlined examples of where using multiple theories of change, enabled tailoring of different inputs for different communities and key populations. Examples included policing in Bougainville, urban safety in Lae and Port Moresby, and targeted interventions for youth employment. These interventions, some of which were externally funded or funded in partnership, seem understood and accounted for at central government level within the MTDP framework. Despite these kinds of positive shifts, line agencies in PNG still reported that their inclusion in planning is limited and this exclusion impacted on their capacity to implement service delivery such as the free primary education policy introduced by GoPNG in 2012, as there were insufficient trained teachers and classrooms available for the students enrolling.

## ***Sustainable development and reporting***

Sustainable development was the second emergent concept evident in the MTDP development process in PNG, which provided some strategic country direction alongside an emphasis on increasing service delivery coverage. The research also noted efforts by GoPNG, through the central planning agency (DNPM), to integrate other dimensions of theory-based evaluative approaches (such as using key sector indicators from MTDP) and performance management (such as collating performance results) at a country and sector level, thereby enabling some progress reporting of PNG sector results.

Portfolio budget approaches aligned to the MTDP sectors also provided some transparency for the allocation of resources and partner inputs towards PNG's development goals. The incorporation of sustainable development at a national level was emphasised as PNG considered how to shift its reliance on oil and gas and expand agriculture and tourism sectors to support a more sustainable economy.

In comparison, in A/NZ underpinning sustainable development themes had limited evidence. While enhancing living standards and wellbeing were emergent concepts, that gaining traction under the current Labour coalition government, it appeared unclear which level of agency (central or line) was accountable for the type and extent of outcomes achieved. There seemed to be a gap evident in coordinating A/NZ's country development across sectors, and it remained unclear which agency would undertake the leadership to plan and coordinate a more sustainable and equitable approach to A/NZ economic and social development.

To improve system and component performance, people need to be aware of bigger world and then see their part in it. We need to know where the money is going and how that's getting down to the people. (A/NZM4)

These findings raised question as to under whose mandate the responsibility to steer A/NZ's strategic direction and provide accountability and sustainability oversight should lie. The current devolved managerial responsibility to A/NZ line agencies seemed to be contributing to fragmented planning and management approaches.

## ***Evaluation***

This third element in the analysis of the *ideas* layer complements the previous two. In A/NZ, there was no mandate for evaluative practice (evidence-based, equitable and transparent) being used systematically in the measurement of results, analysis of need or delivery across regions, sectors and at country level, to coordinate and prioritise funding, or to set local, regional or country goals. Government funding allocation was undertaken by A/NZ Treasury, with each agency submitting budget bids. Participants

considered the prioritisation process was not inclusive, planned or transparent across sectors, regions and even at a country level. This finding highlighted gaps in A/NZ's democratic planning process and adherence to the Treaty of Waitangi principles. Some participants noted that an uncritical maintenance of colonial legacies prevented equitable inclusion and participation for all A/NZ citizens.

PNG's published development plan and central agency mandate for evaluation, legislated for in 2016, provided a more systematic and transparent approach than A/NZ and followed international trends such as the "new" national planning described by Chimhowu et al., (2019, p. 77) which highlighted that more collaborative planning processes were expanding globally. In addition, in PNG provincial planning processes, the findings showed that decentralisation and regional effectiveness as units of analyses might be used alongside evaluation of development effectiveness which may assist in producing more equitable results for citizens. The PNG provincial planning approaches mandated under Organic Law appeared to contribute to provincial administrations liaising with line and central agencies for funding and service delivery implementation. However, PNG's limited accountability processes were recognised constraints.

### ***Non- financial performance information and accountability***

In A/NZ, there was an increased focus on central agencies auditing for non-financial information, but accountability gaps persisted in relation to agency responsibility for the extent of outcomes achieved, and/or substantiating progress at sector and country-level. This meant that interdependent sector and country goals (such as links between coverage and quantity of health services and population needs) did not appear to be regularly or systematically undertaken. This situation differed from developments in PNG. The MTDP, PNG's whole of country plan, was updated at regular intervals and an emerging evaluative use of MTDP data contributed to this process. However, my findings suggested that while the MTDP was increasingly being used as a collaborative, dialogic tool to facilitate development partnerships and align sector inputs from multiple funding partners, PNGs limited auditing and accountability capability undermined their planned intentions.

Given the absence of a published country development plan in A/NZ, it was not surprising that the findings pointed to a need to prioritise sector and agency goals by expanding the use of results frameworks for sectors and programmes. Participants noted that there seemed to be a reluctance by A/NZ agencies to measure and account for the extent of outcomes achieved which may be partly due to the significance of the inequalities displayed. It appeared that changes in development planning and practice were required but the disjunction that was pointed to in this research between the

respective mandates of central and line agencies meant these planning and accountability gaps remain unaddressed. Evaluative work that highlighted disparities is likely to have political effects that can be challenging for the government of the day. The recent report *He matapihi ki te kounga o ngā manaakitanga ā-hauora o Aotearoa: A window on the quality of Aotearoa New Zealand's Health Care* (Health Quality & Safety Commission, 2019) highlighted the strength of systematic evaluation but also illustrated the depth of challenge to “resolve the health inequalities between Māori and non-Māori and advance Māori health” (p. 49).

Participants highlighted that an increased M&E capacity and capability were required in both PNG and A/NZ agency business units and at the corporate level to measure, aggregate results data and report on progress. Increased capability to plan, develop and use result frameworks as well as increased capacity of managers and agency staff to use theory-based evaluative approaches and processes may enable more adaptive approaches. An increase in managerial capability may also assist in meeting an expanding demand for services and interventions addressing A/NZ regional disparities noted in health, education, employment and transport infrastructure. Whereas in PNG, strengthening links and collaboration between DNPM and line agencies were advocated for by research participants to enhance regional service delivery alongside increased support for provincial implementation.

Williams (2015), when discussing impact evaluation, emphasised that a systems approach can assist in addressing “understanding interrelationships, engaging with multiple perspectives and reflecting on boundaries” (p. 8). However, in A/NZ, the function of evaluation at strategic and programme levels remained unclear, which has affected the measurement of development progress. In PNG, a centralised approach to evaluation was under design. In both countries, increased institutionalisation of evaluation, as promoted by Dahler-Larsen and Boodhoo (2019), appeared required to enhance governance, and improve accountability and transparency. Rothstein (cited in Dahler-Larsen et al., 2019) suggested that “a more mature evaluative culture” (p. 283) can support “some normative theories of good governance [that] includes citizen well-being and social equality” (p. 282).

### **8.1.2. Layer 2: Institution**

A key finding highlighted in my analysis was that the focus of the impact model dimensions at the institution level (within this research this contextual layer refers to government line agencies including health, education, justice, transport, social development, economic development, agriculture, and up to two other agencies with a specific cultural or regional focus) was operational and process-based rather than

values-based, contextual and strategic. In this section, I focused on eleven institution limitations or 'pinch-points' that were highlighted by participants and noted that there were also four enabling factors in this layer, which contributed to my theory-building on country approaches to national planning, management and evaluation to enhance development effectiveness. The limitations were grouped under four themes: endogenous management practices, data issues, approaches to organisational learning and exogenous impacts.

### ***Management practices***

Over two-thirds of my A/NZ research participants highlighted that managerial systems in A/NZ agencies were more focused on processes (such as administrative information and project management) than transformative results (such as addressing disparities in health and education outcomes for different population groups, measuring progress and tailoring of inputs). This was important as it meant the issues of inequality and access to tailored services remained. Research participants emphasised that practice could be enhanced by more systematic, inclusive, and collaborative approaches involving customer perspectives from different population groups. Increasing the access and inputs of citizen and customer voices was supported by over half the research participants to enhance service delivery and adapt inputs relevant to changing contexts and peoples' needs such as youth programmes to enable more sustainable outcomes.

The separation of corporate functions (such as planning, performance, reporting and financial) from M&E focused on activities and services further increased emphasis on operations rather than service value. The fragmented agency approaches and limited collaboration within agencies by different teams (such as strategy, policy, reporting and communications) were noted as contributing to silos and the duplication of functions. My research highlighted there was a need to integrate the operational planning and management functions more directly with measuring and reporting of service delivery results. As an evaluation practitioner, I considered such a shift in focus could help to align functions and reduce duplication of functions within agencies and may also improve accountability within agencies and transparency of resource choice and allocation. However, this will require a very clear institutional narrative and potential reconfiguration of the role and nature of strategic evaluation within management, or a new paradigm.

Furthermore, the separation of corporate functions in A/NZ (such that strategy, performance and reporting were separated from evaluation functions with the latter positioned as an operational activity) may either be unnecessarily adding to agency complexity or overlooking the value of evaluative approaches at the strategic level. The research suggested these functions could be streamlined, and relationships empowered

by involving stakeholders using embedded evaluative approaches (including qualitative and quantitative methodology) underpinned by values of inclusion, partnership and participation. Stickl, Haugen and Chouinard (2018) supported this approach supported in their findings on relational power which emphasised that “cultural responsiveness requires active stakeholder engagement and participation” (cited in Chouinard & Cousins, 2009, p. 383). Research participants considered that aligned and integrated evaluative approaches might enable more inclusive sustainable development that supports citizens and their democratic rights for more equitable access to services.

Throughout the organisation getting really clear about the purpose of the work that we’re doing, what the problem is that we’re seeking to solve and what we’re seeking to achieve, so getting clear about things like outcomes and benefits. It’s harder to come up with systems that are simple and meaningful for running the business and systems that make sense to everybody even the sort of the lowliest person working on a project who understands why they have to do reporting every month. That’s kind of where you need to get to. (A/NZM8)

My analysis also suggested that due to A/NZ’s devolved managerial approach to agencies, priority areas may be buried within each agency. This could be why regional inequalities and unmet needs of ethnic and priority groups were disproportionately represented in health and education. Moreover, the research evidence pointed to a remaining functional separation of policy and performance teams within A/NZ agencies which may be leading to information gaps. A/NZ participants considered both structural reform (such as a mandate for more centralised planning and evaluation) between central and line agencies alongside the integration of functions could provide more clarity for accountabilities and increased flexibility to adapt for different contexts and needs.

There’s a big role for central agencies to try and keep that focus on long-term outcomes and to help, because a lot of these things are using value for money data on long-term outcomes, thinking of performance management as tracking progress over time, remaining sustainable, all these things are mutually reinforcing and all part of a package. (A/NZM20)

In both PNG and A/NZ, research participants highlighted that shifting the focus of leadership from chief executives to horizontal management teams and vertically to regions was desirable. New approaches were required, including “vision, evidence and capacities at all levels of government” (OECD, 2017, p. 37). A/NZ participants emphasised the need for chief executives and management teams to be strategically driven and manage focus on outcomes as opposed to processes. This shift of emphasis was a key difference of focus from the dimensions in the impact models (initial and

updated) which focused on management processes and the underpinning assumption that development results would be enhanced through improved management processes separated from M&E processes.

### ***Data issues***

In both PNG and A/NZ, the limited availability and selective use of relevant data may hinder measurement of progress and assessment of development effectiveness. Data governance and ownership were key issues raised by research participants. In addition, it appeared that limited data was aggregated from programmes and services to a sector level in both PNG and A/NZ. Therefore, increased emphasis on collection and use of segmented service and intervention data for ethnicity age and gender, and aggregation of data was wanted and needed from multiple regional locations in both PNG and A/NZ. My research indicated that limited relevant and available data was impacting negatively on the transparency of and accountability for development results. This lack of outcome data remained unaddressed in A/NZ and PNG.

The use of data in A/NZ agencies was compliance-based, rather than being undertaken through an organisational learning lens with an adaptive management intent required and measured at regular intervals.

Without that longitudinal view data is quite meaningless but if that's natural in time data that the formal system demands, and so then if you have your finance team and your external reporting people focused on generating that then they have a good understanding of what they need is longitudinal data but they're not giving you space to focus on that. (A/NZE3)

The evidence suggested sector-wide approaches tended to have a more considered use of data and analyses. In addition, research participants in both PNG and A/NZ suggested that use of data in decision-making remained relatively limited. Further attention is required through increased capability (amongst managers and leadership teams) and system design (such as increased use of results frameworks across line agency programmes and services, data collection IT systems and aggregated results reporting). The linking of results frameworks or logic models and systems concepts was promoted (Renger, 2015) as situating the results framework/ logic model within a broader system model context (Renger, Atkinson, Renger, et al., 2019) assists to understand interrelationships and system components, which were key concepts that contributed to my theory building.

Research participants also pointed to the need for increased provincial, provider and community inclusion and participation, which may assist with data collection and

reporting of progress in both PNG and A/NZ regions and by service providers providing relevant data to agencies.

### ***Organisational learning***

Proactive interpersonal relations and sharing of information, particularly between line and central agencies and their wider regional stakeholders were highlighted as areas requiring increased attention in both PNG and A/NZ. The findings also suggested that such internal organisational culture has an impact on the transparency of results. Participants felt that managers need to be more open to learning from implementation and feedback. Research participants considered that people could learn and adapt to updated approaches as advocated by Agryris (1997). They noted that risk-averse organisational cultures provided limited support for internal learning and made it more difficult to adopt more systematic approaches to planning and implementation.

Being a learning organisation takes time and good practice to embed learning across organisations. Managers at different levels and offices need to have higher capability, and there is slim resourcing compared to other teams within agencies such as policy. Accountabilities need to be clarified of who is managing and who's accountable. (A/NZM4)

The analysis also highlighted a limited use of systematic approaches by line and central agencies in both PNG and A/NZ for reflective learning and timely feedback. As an evaluation practitioner, I considered that more explicit guidance and increased central agency governance and accountability oversight may be required.

### ***External factors***

A/NZ agency participants reported that the three-year election term in A/NZ impacted on planning and progress. Findings pointed to tension between A/NZ ministers' short-term focus and longer-term objectives. These were problems in areas such as health and transport planning where longer time frames were deemed critical. The short-term focus appeared to hinder A/NZ agencies in developing more responsive and adaptive learning approaches and performance cultures with evidenced-based feedback, which in turn may be constraining development results.

The best place to basically locate some sort of accountability for getting the focus on outcomes must be in the delivery and the operational part of the executive not the political part of the executive. So, in terms of demonstrating real change, I think we still have some way to go in terms of being able to demonstration attribution. (A/NZM4)



In PNG, the mandated country-level planning framework pointed to a more systematic approach to linking sector goals and allocation of funding beyond parliamentary cycles. An increasing focus on country level design for evaluative approaches and methodology in PNG seemed to be supporting a more systematic M&E approach across sectors within DNPM. DPLGA supported a focus on agency and provincial administration leadership capability.

The focus is on the people mind-set and leadership. We've started, you can have a good organisation framework, but its people that make it work, and also leadership you need a good leader. And you need people with a good mind-set to accept change. And we're dealing with human beings we have emotions, feelings, everybody has different mind-set. So, we need to work along to really change that mind-set to be able to work as a team. And work as a body to implement services. I think only then we can achieve. (PNGM10)

It was at this point in my theory-building reflections as an evaluation practitioner that I was considered how evaluation at the strategic level could actually work. I realised that the nature and role of monitoring and evaluation would need to be raised up to the wider infrastructure layer with clear guidance and direction from central agencies. My research findings showed that it was not undertaken systematically if left to each line agency (as in A/NZ) or to regions (as in PNG) to design and implement approaches.

### ***Enabling factors***

Participants also noted a number of insights that suggest positive developments were more strategic, contextual and values-based. Using systems approaches was promoted by both managers and evaluators as it encouraged the integrating and embedding of more evaluative conceptual and operational approaches. In A/NZ, some agencies have shifted to more sector-based approaches such as the social development and justice sectors. Regional offices were increasingly providing information that could be used to tailor services for different peoples' needs, and these offices may be useful in determining system boundaries in a regional context.

Embedded adaptive processes such as those used by MFAT appeared to enhance a strategic development focus reliant on more integrated planning and evaluation functions. Embedded processes were noted as reducing duplication, keeping managers focused on the strategic intent and including stakeholders in feedback loops, supporting learning and decision-making.

However, in A/NZ as there was no mandate for evaluation, these more integrated and adaptive approaches were not systematically used. In addition, managerial constraints (such as leadership confidence and coordination skills) at PNG provincial government

level were indicated as likely to be affecting service delivery planning and implementation.

An emerging agency structure gaining traction in both PNG and A/NZ. Agencies involved the use of thematic groupings of business units around specific service delivery focus areas (such as community development). Such an approach was supporting more integrated operational approaches. However, despite repeated restructuring in A/NZ agencies, the separation of corporate management functions for planning, performance and reporting remains. Limited evaluation is undertaken, particularly at the strategic level, hampering agency learning and performance.

Finally, the findings highlighted that increasing regional inclusion, participation and governance under development partnerships may enable co-design and adaptation to tailor services and interventions for specific population groups. Such approaches may be enhanced through increased collection and use of segmented service and intervention data enabled by using centralised IT systems for data collection accessed from multiple locations. Such a shift may enable managers to adapt inputs in collaborative partnerships. This might involve working together to decide what to do more or less of depending on what is needed or using cultural and age-appropriate approaches to enhance service delivery and interventions. In addition, these partnership approaches may enhance more relevant and timely data collection enabled by advances in technological functionality. This could facilitate more regional data collection and greater engagement with more “deliberative processes” (Dobell, 2003, p. 7). An applied use of digital platforms was also supported by Dormer and Ward (2018) for assisting with integrating institution functions (such as performance reporting and data collection) and “increasing citizen engagement” (p. 32).

Overall, in terms of the institution level, the findings pointed to a need for greater integration and streamlining of organisational processes underpinned by more adaptive and streamlined results-focused evaluative approaches. Any transformative shift in focus towards results rather than processes will be dependent on more inclusive, participatory and democratic approaches that entail partnerships predicated on regional, provider and community inclusion.

The importance of understanding context, relationships and the needs of priority groups were highlighted by research participants in both PNG and A/NZ. Approaches needed to be more inclusive and undertaken authentically, hence the promotion of a more values-based management approach with underpinning values of inclusion, partnership and participation.

### **8.1.3. Layer 3: Wider infrastructure**

The context of power and political organisation under which both institutional level change and the uptake of conceptual ideas and frameworks were facilitated (or not) comprises this third analytical section. Howlett (2002), when discussing administrative styles, noted that “rules, norms, and symbols affect political behaviour; that the organisation of governmental institutions affect what the state does; and that unique patterns of historical development constrain future choices” (p. 27). A/NZ and PNG operated under different political conditions despite both being Westminster-style democracies.

What particularly characterised this difference in terms of the ‘development’ ethos was that A/NZ opts for a ‘freedom to manage’ approach in line with its neoliberal orientation and PNG used a more centralised planning and control orientation that aligns with its development planning approach was used to support PNG’s national unity given its diverse geography and ethnic make-up. As Castles (2001, cited in Howett, 2002) noted, “distinct national administrative cultures have an impact on national policy outcomes, and nations tend to follow the precepts of the administrative models from which they emerged” (p. 18). These kinds of differences were identified by participants and were discussed below.

#### ***Development approaches***

A/NZ participants pointed to a gap in development capacity. Agency managers appeared process-focused, and operations segmented across strategy, performance, reporting, finance and policy. Evaluation was positioned at a predominantly operational level and used in different forms in all or any of these organisational segments in predominantly transactional and quality assurance roles. There was limited evidence of development-focused planning and no evidence of evaluation used as a strategic development tool in government agencies. There was an emerging focus on wellbeing and living standards (an overall strategic focus) at central agency level driven by the Treasury, and some thematic approaches were used by sector-based agencies such as justice and social development. However, participants noted the lack of tools and approaches to respond effectively to this political reorientation.

By way of contrast, the mandated country planning frameworks adopted in PNG, pointed to a more considered development approach that more directly linked sector portfolios with the allocation of funding. An increased focus on country level design for evaluative approaches and methodology was displayed in PNG. The government mandated central agency functions appeared to be supporting the development of a more systematic M&E approach across sectors with links to agencies and provincial

administrations. This was enabling PNG to report at country level on key performance indicators across multiple sectors which contributed some evidenced-based data to the annual portfolio-based budget allocation processes.

### ***High-level guidance***

Research participants suggested that what appears to be a 'hesitation' by A/NZ central agencies to provide more operational guidance may be a response to the more prescriptive centralised government control introduced in the 1980s. However, my research indicated that agencies were, in fact, looking for more prescriptive, systematic approaches (with some flexibility), particularly for integrating and implementing evaluative functions coordinated by central agencies. Both managers and evaluators endorsed such integrated, high-level guidance approaches in PNG and A/NZ. This might enable greater collaboration between agencies and ensuring inclusion and participation occurs more systematically between sectors and regions. The need for more prescriptive guidance and support with accountability oversight by central agencies in the A/NZ case was to some extent highlighted by the PNG example of more centralised design and coordination of M&E approaches. The more coordinated PNG approach mirrored an expanding trend in national development globally of increasing roles being undertaken by central agencies (Chimhowu et al., 2019).

PNG's use of a country plan with a growing emphasis on sustainable development was regarded as providing some implementation continuity from central to line agencies. This finding may highlight options for A/NZ central agencies to consider. The question of whether development of a high-level country plan would enhance country and sector direction underpinned by values of inclusion, partnership and participation was yet to be explored in the A/NZ context. My research suggested the possibility that such a turn might enhance and transform development results and governance.

My analysis suggested that PNG's country planning mandate and framework supported a systematic approach to sector planning and enabled a portfolio-based process for funding allocation. In addition, PNG's planning framework enabled a systematic inclusion of provincial-level concerns into country-level prioritisation that was more cognisant of geographical constraints and tribal cultures. In contrast, A/NZ's devolved management approach to agency chief executives with the freedom to manage appeared to produce duplicative, or ambiguous administrative systems and processes.

When analysed against Howlett's (2002) *Multi-layered Concept of an Administrative Styles* for regional inclusion and prioritisation of inputs, for example, it falls short against the level of state participation and may be limiting access to services. My evidence suggested that some regional processes were in place in A/NZ through regional councils

– such as for transport planning inputs and DHBs – and that regional coordination hubs were emerging in the justice and education sectors. In the absence of an A/NZ country plan, however, how prioritisation occurred at a country level across sectors was not transparent. Research participants considered budget allocation by A/NZ Treasury lacked transparency and was overly influenced by the three-year election term.

Furthermore, regional administrative boundaries for planning and service delivery were not congruent. The plethora of regional administration units in A/NZ government levels (central, regional and local) was a recognised issue (Gray, 2002) which contributed to fragmenting services. My research indicated that regional structural reform might be required in A/NZ to streamline processes and systems for sector and agency regional service delivery, combined with an increased national planning mandate and oversight by central agencies.

In addition, my analysis pointed to administrative and accountability gaps in A/NZ, and what I considered could be viewed as ‘unproductive swirl’ generated by often rapidly changing managers attempting to develop policy on an ad hoc basis, and re-prioritise activities or working to achieve results with limited strategic country and sector direction.

Equity concerns in accessing services and interventions were highlighted by research participants in A/NZ, particularly for priority groups such as Māori, Pacific peoples, and youth. These issues raised questions on the role of the state in A/NZ for supporting inclusion and access to services throughout A/NZ. Evidence from the participants in my research suggested this situation has yet to be addressed in A/NZ.

A/NZ agencies used an agency-based planning and decision-making process in conjunction with A/NZ’s three-year political cycle. These agency-based development approaches and narratives were a legacy of A/NZ’s history of moving from a welfare focus to more devolved managerial, neoliberal free-market (Chang, 2003, p. 4) and social investment approaches. The more current focus on wellbeing and living standards may have derived some of its origins in the 2015 global endorsement to the 2030 SDG Agenda. A/NZ was a signatory to this 2014 UN resolution (UN, 2015), which included countries considered both developed and developing. Research participants highlighted that it remained unclear to them how wellbeing (with some documentation evident linking to SDGs) becomes embedded across A/NZ’s state sector as a development narrative and implemented.

I almost think it is an issue of culture change; unfortunately, this is a bit fluffy and amorphous, but I think too many agencies focus on short-term goals. It is almost if you had a big international business and all your branches are doing their own thing, and you need a big shakeup from corporate heads to say no this is our new

goal, everybody is going to be working towards this and keep the concerted focus on that. (A/NZM20)

### ***Service delivery - access and inclusion***

My analysis also pointed to tensions being evident in power, control and inclusion between agency national offices and devolution to regions. Evidence from A/NZ and PNG participants suggested that tailored service delivery with localised feedback loops would be advantageous for service consumers. A/NZ participants noted that the customer or beneficiary voices were not included enough in planning and feedback processes. A/NZ participants emphasised that data systems needed to be tailored and made available to regional offices to focus understanding of development progress at regional levels. In A/NZ, this regional devolution has existed in relation to environmental and transport policy for some time, and it was also underway for social development. This finding supported Grey's view of the need to improve A/NZ's regional alignment and sector coordination. Also, A/NZ line agency national offices and central agencies appeared to remain mostly involved in transactional processes which were underpinned by some country-level forecasting from A/NZ Treasury. My research highlighted that this gap in feedback for country-level planning and coordination across sectors at national and regional levels in A/NZ remained unaddressed.

PNG appeared to have more systematic inclusion and coordination for service delivery through its mandated provincial administration and cascading levels of local government. Challenges in service delivery coverage and sustainability of services were, however, highlighted by research participants. Both sustainability and coverage constraints were recognised by central agencies such as DNPM and DPLGA, with noted efforts underway to build provincial leadership and managerial capability and capacity. It appeared that links between PNG's central agencies, line agencies and out to provincial government administrations required strengthening to support provincial priorities and needs. A recognised issue for PNG was the limited accountability of funding flows which was compounded by what seems to be a lack of political will to expand audit at a central level.

Issues in equity and inclusion in accessing services and interventions, particularly for priority groups such as Māori, Pacific and youth, were highlighted by research participants. The gaps in overarching planning and direction suggested an increased role for central state planning in A/NZ may be required to support inclusion and access more successfully to services throughout A/NZ. As Gray (2002) suggested:

Despite calls for a 'whole of government' approach, cross-fertilisation between the sectors appears to be limited. In New Zealand, the potential for regional co-

ordination is complicated by the lack of alignment of regional boundaries both within and between government departments, and between central and local government and other service agencies. Iwi boundaries add to the complexity. (p. 48)

In both PNG and A/NZ, questions relating to the role of the state in balancing capitalism, free market and interventionist approaches for systematic inclusion and democratic access to services were largely unaddressed in current research. These themes highlighted for me in my theory-building that more direct guidance and increased governance oversight with an equitable accountability lens may be required.

### ***Population equity and inclusion***

PNG with its country plan appeared to have a more aligned and explicit development narrative congruent with the 1975 constitution. It promoted sharing of wealth, a focus on wellbeing and recognition of PNGs diverse tribes and geography. The mandated planning framework operationalised the constitutional intent which provided a basis for priority setting dialogue across PNG's provinces.

In A/NZ, under the terms of the Treaty of Waitangi, principles of engagement and partnership have been articulated, yet the research highlighted it was still unclear how these principles might be implemented. Individual agencies were left alone to determine how to operationalise Treaty obligations.

Thus, a key finding of this research was that the inclusion of key population groups (Māori and Pacific) was not occurring systematically in A/NZ. Research participants referred readily to relevant legislation that governs the terms of engagement, such as the Auditor-General Accounting Standard (AG4) the Public Finance Act (2004) and the Treaty. Yet, there appeared to be an absence of proactive longer-term strategic goals and planning with annual targets identified to achieve these goals and clarity of how progress will be measured. In A/NZ, there appeared to be a limited country strategic direction and mandate for coordination between central and line agencies which research participants considered contributed to these strategic planning and accountability gaps. Given these gaps and a government-wide focus on processes such as business case preparation and project management, my findings highlighted how A/NZ line agencies had evolved separate managerial focused processes to operationalise their devolved responsibilities. This research highlighted gaps and potential opportunities in A/NZ for evolving development, management and evaluation approaches given its recognised low level of regulations and laws in this interdisciplinary interface.

I think there's a big role for central agencies to try and keep that focus on long-term outcomes and to help, they're using value for money data, keeping your eyes on a long-term outcome, thinking of performance management as tracking progress over time, remaining sustainable, I think all these things are mutually reinforcing and they're all part of a package. (A/NZM8)

In addition, participants in both countries noted that the availability of segmented development data to aggregate and assess overall development effectiveness was limited. This finding suggested that central agencies may need to focus more on aggregating and analysing evidence in relation to development priorities to assess progress and equity. Research participants considered both central government and line agencies could shift their focus from transactional processes to become more transformative by focusing more on achieving and measuring development results and improving lives for citizens. This may require more explicit central agency guidance and governance oversight.

### ***Sustainable development***

In A/NZ, the triple sustainable development dimensions (social, economic and environment) did not appear as an active combination of concepts underpinning development. Whereas, PNG has actively developed a sustainable development country strategy and appeared to be embedding these triple SDG dimensions in successive country plans.

For sustainable, inclusive development to become active rather than just an articulated goal, citizens' voices and democratic rights may also need to be stepped up or mandated. To ensure more inclusive participation in development partnerships that are authentic in nature (as highlighted in this research) and development and managerial approaches, A/NZ may require structural reform with a more significant role undertaken by central agencies. An expanded central agency role appeared underway in PNG where DNPM as the national planning agency was displaying efforts to utilise a more centralised approach for embedding sustainable development and monitoring and evaluation. This expanded central agency role for DNPM was enabling planning for and measuring progress within its country and sector development goals, budget allocation and reporting. However, PNGs recognised accountability issues appeared to remain, which undermines PNG's development progress.

#### **8.1.4. Summary**

Both PNG and A/NZ may need to consider the kinds of factors that emerged under the *wider infrastructure* category in the analysis. The kinds of questions that realist Pawson (2006) asked were relevant in the context of my overall argument for a values-



based evaluative management approach. Pawson asked: “does the intervention have the political backing to drive it .... Are there resources to underpin it? ... Is there public support? (p. 31). PNG’s provincial government structure and planning mandate appeared to provide for a systematic inclusion of provincial priorities which supported service delivery coordination in the provinces, with embedded central government approaches for longer-term planning, portfolio budget allocation sustainable development and evaluation, providing a mandate beyond the political will to drive these priorities were seen to be present.

In A/NZ, these kinds of considerations were less clearly underpinned by political will or when the will manifests itself under one government, and it may change under the next. A/NZs expanded accountability approaches for financial and increasingly non-financial information, provided a clearer pathway for resource flows, and such an approach could improve PNG accountability and development results data. Inclusion and participation by providers and communities in planning, managing and providing feedback in policy design and processes were areas highlighted from this research for increased attention and explicit consideration that may enhance development effectiveness and governance.

## **8.2 Potential reconfiguring for the role of strategic evaluation**

The research undertaken for this study developed evidence that pointed to several shifts that may need to occur within government institutions to enhance development effectiveness and governance. The analysis of findings in Chapters Five, Six and Seven suggested there was a need to use a more values-based management approach (that was focused on inclusivity, participation and partnership). They also suggested that corporate management functions should be more integrated with strategic level evaluation.

I argued that the current systems of management and governance seem to fall short of SDG goals in several ways which since 2015 have grown in global importance. First, inclusion of key population groups and regions into development planning, implementation and assessing progress was not systematically occurring which was impacting on equitable access to relevant services. Second, the current management approaches did not ensure that evidenced-based data is available for resource choice and decision-making. Third, currently, the progress and sustainability of social, economic and environmental outcomes were not measured or assessed systematically across sectors. All of these are a critical foundation for working towards country and global SDGs. Given that countries have committed to the Sustainable Development agenda in

2015, it was plausible to assume they also take responsibility to ensure contribution to those goals is effective and transparent. This meant that each country has a responsibility to ensure an evidence-based and informed basis for development is undertaken as part of its country governance. However, in 2018 A/NZ MFAT produced the first sustainable development reporting against SDGs, yet it appeared unclear how this report related to key A/NZ agency management activities and reporting. Moreover, PNG has a country constitution to adhere to, and A/NZ has the Treaty of Waitangi recognised under international law underpinning country development.

The research findings and implications to be drawn from them are significant as they challenge both A/NZ's devolved managerial mandate to line agencies and PNG's funding to line agencies for regional service delivery. They also suggested that more explicit consideration might be required for the nature, roles and approaches of the state, levels of government, public sector management and development partners. If the articulated goal was inclusive sustainable development, then the current paradigm of development, management and evaluation constrained inclusion, governance, accountability, and sustainable development.

As a researcher and evaluator, I found these research findings confronting as they challenged the current positioning of evaluation in national development contexts as a transactional activity at the operational level. What I take from this theory-building research, however, was that there was some potential or demand to reposition evaluation such that it could be more widely used to support strategic management.

I am, therefore, proposing a paradigm shift for evaluation that could be specified as "evaluative management". The findings led me to two insights: first, that there were additional key concepts at the interface of development, management and evaluation evident beyond those currently specified in development literature; and second, that a re-positioning and integration of management and evaluation approaches, concepts, functions and roles could enhance development effectiveness and governance. The research participants identified that the concepts characterised as emergent were those that are required to enhance development effectiveness and governance.

My theory-building proposal for further enhancement of development effectiveness and governance was that central agency approaches could further support increased inclusion and substantiation of development results through feedback loops and adaptive processes. One alternative was to configure a role for strategic evaluation that extended my country system typology with three contextual layers and employed the kinds of concepts listed in Table 14. Such an approach may require an overarching centralised design (for planning and strategic evaluation) with central and line agency oversight. This

may potentially require further mandates to support the inclusion of key population groups and regional priorities in planning and evaluating the progress towards their development, and more localised service delivery coordination. This would mean for A/NZ that a mandate for evaluation is made (as PNG has undertaken) to centrally plan, measure, and report on country and sector goals and progress, and align with A/NZ Treasury to allocate the budget using a more transparent and evidenced-based prioritisation process underpinned by more inclusive sustainable development values and principles. This may, in turn, assist in addressing inequalities for population groups and regions and contribute to more sustainable development.

This theory-building research also pointed to a potential reconfiguring of strategic evaluation (with associated concepts listed in Table 14) that is enabled under a 'new' paradigm of an 'evaluative management' approach within more aligned central and corporate agency functions together with line agency integration of corporate functions (including planning, performance management, business processes and reporting). This reconfiguring and integration may enable more inclusive sustainable development and governance.

### **8.3 Identifiable concepts requiring more attention and inclusion**

The final element for this thesis related to good practice concepts. My country system typology with three contextual layers was useful to analyse concepts evident from my findings. I was able to distil some knowledge, and practice concepts for managers and evaluators to more explicitly consider that potentially could enhance development effectiveness and governance in different contexts. The literature, interviews and government documents all contributed ideas and information about what constituted 'good practice', and in this next section, I focused on these concepts to provide further insight by using critical realism strata (Wuisman, 2005) of experiences, events and mechanisms underpinning these concepts. The emphasis in this part of the discussion was on those concepts that pointed to where changes in practice and direction might need to be considered.

During my theory-building research analysis, I considered that the *ideas, interventions and frameworks* layer needed to be conceptually re-positioned beneath the *wider infrastructure* layer as a new form of 'strategic evaluation' within a new paradigm of *evaluative management*. The rationale for this repositioning was that I considered the operationally focused positioning of evaluation as shown in my literature review could exclude and prevent evaluative approaches and insights structurally and strategically from contributing to strategic management and development discourse and practices

that my analysis showed required attention and inclusion. I argued that without a planning and evaluation mandates, such as in PNG, integrated development planning, management and evaluation may not be undertaken systematically and equitably for population segments and regions, which impacts on citizens' democratic rights.

I grouped concepts as summarised in Table 14. A discussion of these fourteen concepts, referenced to findings that also appeared in the literature, grouped across the three contextual layers were provided in detail in the following section.

**Table 14. Key identifiable good practice concepts requiring further attention or inclusion**

Contextual layer	Concepts requiring more attention and inclusion
Wider infrastructure (Central agencies and global trends)	<ol style="list-style-type: none"> <li>1. Deliberative governance</li> <li>2. Centralisation and devolution</li> <li>3. Development approach incorporating values of inclusion, partnership, and participation</li> <li>4. Leadership, management, power, and control</li> <li>5. Regionalisation and segmentation</li> </ol>
Ideas, interventions, and frameworks (reconfigured as Strategic Evaluation)	<ol style="list-style-type: none"> <li>6. Strategic intent, context, and systems</li> <li>7. Inclusion, partnership, and participation</li> <li>8. Integrating strategy, planning, theory based evaluative approaches, and performance management</li> <li>9. Integrating adaptive management approaches and portfolio-based funding</li> <li>10. Management and evaluation capability</li> <li>11. Community inclusion, empowerment, and wellbeing</li> </ol>
Institution (line agencies)	<ol style="list-style-type: none"> <li>12. Integration of development, management, evaluation, and technology functions</li> <li>13. Enabling interpersonal organisational culture for feedback, learning and performance, decision-making</li> <li>14. Mixed-methodology, data and evidence, analytical insights, and integrated reporting</li> </ol>

### **8.3.1. Wider infrastructure layer**

In the *wider infrastructure* layer, I grouped five concepts: governance and democracy; centralisation and devolution; values relating to inclusion, partnership and participation, development effectiveness and sustainability (including wellbeing and resilience); leadership, management, power and control; and regionalisation and segmentation. These were discussed below to highlight the relevance of these concepts for enhancing development effectiveness and governance, thus contributing to more inclusive sustainable development.

#### ***Deliberative governance***

First, I considered that a concept of deliberative governance could overarch country, sector, regional and programme development contexts where underpinning ideology

choices were made such as “deliberative democracy” which “demands not only participation but also equal access to decisions by citizens with a stake in them” (Davies, 2007, p. 781) building on work by Dryzek (2000), Medearis (2005) and White (2003). One PNG manager emphasised that “development effectiveness is for all citizens” (PNGM10) highlighting that considered democratic decisions by managers were needed on who was included, and how development was undertaken. My research showed that more deliberate governance oversight might be required to ensure more systematic line agencies practices that support inclusion and participation.

However, my research highlighted that current country and agency approaches, particularly in A/NZ did not appear to be deliberate or democratic. Significant population groups appeared under-represented in planning and progress assessment. Five out of six key A/NZ government agencies were not reporting segmented data which can bury actual results. Increased direction and an oversight mandate with more operational guidance may be required to improve governance and management capability. This approach concurred with Grindle’s (2004) “four capacities to state” advocated by an A/NZ research participant (A/NZM16) to enhance state and managerial capability, and governance. PNG’s country plan, mandated cascading planning framework and legislation such as Organic Law appeared to provide a basis for more transparent and systematic approaches in development and management. This included evaluative approaches which were increasingly evolving and becoming embedded through DNPM, PNG’s centralised planning agency.

### ***Centralisation and devolution***

My research highlighted that balance was required between centralisation and devolution of strategy and management processes within national development contexts to enable regions to undertake more systematic, inclusive and participatory partnership approaches with regional government administrators, NGOs, tribes and private sector. Participants advocated for an increased role by central agencies in setting country and sector strategic direction and integration of strategic evaluative approaches to enable more coherent approaches and processes which could then cascade down into corporate management functions within agencies and regions.

This research finding directly challenged A/NZ’s devolved managerial approach from central agencies to line agencies which appeared to focus predominantly on transactional processes. A rationale for an increased strategic direction by central agencies underpinned by values and functional integration was that I considered it is at the country level where systematic inclusive approaches may need to be mandated

within development planning. This could support more systematic inclusive sustainable development and 'deliberate and democratic' governance.

Eyben (2008) promoted using a more relational approach to aid to deal with power, mutual accountability and responsibility and advocated that "decentralising decision making to as low a level as possible seemed an obvious step to embracing complexity" (p. 48). Yet, my research highlighted that if centralised directive and oversight roles were not provided, ethnic, regional and community inclusion may not occur systematically, thus excluding perspectives of 'citizens with a stake' in development planning and their iterative inputs into implementation and governance.

### ***Development approach based on values***

This research was undertaken between 2010 and 2016, and during this period a relatively unrecognised shift took place as countries (considered both 'developed' and 'developing') completed or evolved their national development plans. Chimhowu et al. (2019) referred to this shift as a trend towards "new national development planning" where the number of development plans increased from "about 62 in 2006 to 134 in 2018" (p. 76) covering over 80 percent of the world's population. The shift also encompassed more practice-led "collaborative, communicative and socially embedded" approaches which differed from earlier more technical planning approaches.

My research found that GoPNG was increasingly using their development plan to align funders and build a more collaborative approach within sectors. In the A/NZ context research participants advocated for more inclusive and participatory approaches as the separate agency-based approaches can fragment stakeholders due to a predominant focus on agency processes rather than the customers. This can lead to the exclusion of key population groups participating in collaborative dialogue and feedback. Both PNG and A/NZ participants wanted more 'authentic' partnership and participatory approaches and processes used.

A key finding from this research was that values such as inclusion, partnership and participation could underpin more integrated approaches of development, management and evaluation. This may assist in aligning managerial functions, promote collaboration and increase social inclusion.

### ***Leadership, power and control***

The research highlighted that there appeared to be leadership capability gaps at central, agency and regional levels. Research participants emphasised incorporating more of a business model focus on outcomes as opposed to processes. Participants promoted expanding leadership and management functions horizontally to include

management teams from different functional areas and at different vertical levels (such as country, sector, regional), alongside enhancing skills and knowledge in evaluative and adaptive approaches. Innes and Booher (2018) emphasised that a “great strength of a collaborative process is its ability to mirror and adjust to change” (p. 90). Using collaborative management and leadership approaches may enable more ‘reflexive, interpretative and real’ (Alversson & Skoldberg, 2009, p. 41) feedback from wider stakeholders that can be responded to iteratively.

Integrating planning, implementation, evaluation and reporting appeared required, and management capability could be expanded through training and technical support. Extending the use of integrated IT platforms provides an opportunity to integrate functions and increase customer inputs (Dobell, 2003; Dormer & Ward, 2018)

In A/NZ, central agency focus appeared to be predominately on chief executive capability development influenced by Sen’s (1989) capabilities, as opposed to a more diverse management team or vertical team focus to enhance sector and regional capability. In comparison, GoPNG appeared to have recognised that more comprehensive leadership capability was required to enhance development effectiveness and governance with an increased focus on provincial leadership capability and capacity development.

Themes of power and control were evident in my research, particularly within A/NZ agencies. There appeared to be limited inclusion of communities and service providers by agency personnel undertaking planning, assessing or progress evaluation. This seemed to hinder a use of adaptive approaches and affects relevancy and timeliness of data collection and feedback loops. Pieterse (2010) in his discussion on power and control suggested using “culture as a device to nation-building” (p. 65) to assist in overcoming colonial legacies. He emphasised that a “national culture will be developed by and emerge from the real people” (p. 65). However, A/NZ appeared to persist with its predominant colonial narrative in which the Treaty of Waitangi was only upheld operationally under case law. This research pointed to ongoing issues of power and control for Māori and Pacific people by A/NZ Europeans. My research highlighted that approaches for inclusion of key population groups such as Māori and Pacific people were not happening systematically. Dalziel and Saunders (2014) argued that “the Māori text of the Treaty of Waitangi provided a powerful expression of this framework [for respecting citizen’s rights] that can serve as a model for all groups of citizens in our relations with central government” (p. 10). My research pointed to a solution of central agencies undertaking an increased role in A/NZ’s strategic direction, providing more systematic integrated development, management and evaluative approaches to agencies which could be cascaded out to regions with a deliberative governance oversight function to

support more systematic inclusion and democratic representation of A/NZ's citizens. I considered A/NZ could undertake these shifts as part of its approach within the 2030 agenda for global SGDs. However, these issues remain unaddressed.

Evidence of increased ownership of their country's development by GoPNG was displayed in the expanded use of their country plan to align inputs towards PNG's development goals with their multiple development partners including provinces with different tribes. PNG appeared to be using its country plan as tool to assist development as a nation particularly given its inherent tribal nature with provincial inclusion recognised under by Organic Law.

### ***Regionalisation and segmentation***

PNG research participants highlighted concepts of 'decentralisation and regional effectiveness'. These concepts were promoted as dimensions of development effectiveness where analyses were broken down by region using segmented data for population, gender and age groups and progress assessed against service delivery coverage and performance. I considered these analyses could assist in addressing development inequalities at regional levels and provide central and line agencies with more relevant data for decision-making. In addition, an increased regional focus could enable timely and more considered inputs for regions tailored to their development. This view was supported by Innes and Booher (2018) who advocated for "genuine dialogue" (p. 18) and participation of communities in their development.

The PNG mandate for provincial governments under Organic Law appeared to support regional priorities for services and interventions and some citizen input. A/NZ's mixed administrative boundaries for different sectors such as health, transport and education contributed to further complexity for agencies. In addition, A/NZ's unaligned regional administrative structures may be hindering citizen inclusion, partnerships and participation with line agencies, private sector and iwi (Māori tribes). This was an area I identified for further research.

### **8.3.2. Ideas layer reconfigured as Strategic Evaluation**

The following section highlighted six conceptual groupings that characterised the space I designated as *strategic evaluation* within an *evaluative management* approach. This designation pointed to my emerging argument that a new space for evaluation needed to be carved out within management that represented an alternative to current practices.

I also considered that A/NZ's shift to more inclusive sustainable development remained constrained without having a designated *evaluative management* approach



and potential *strategic evaluation* function, setting direction at central agency level with guidance and oversight provided for line agencies, and potentially a mandate.

### ***Strategic intent, systems and contextual sensitivity***

The first concept grouping involved clarifying strategic intent and using systems approaches that supported line agencies moving towards a more business-focused direction such as achieving outcomes. These findings aligned with Mintzberg and Water's (1985) earlier work on "emergent and deliberate strategy" (cited in McEwan, 2016, p. 11). My analysis suggested a shift by line agencies towards strategic intent as endorsed by Quesnel (2009) rather than policy and process focus, may enable managers to place development results at the forefront. Such a shift in focus could then be underpinned with more inclusive approaches including increased participation by 'citizens with a stake' in setting direction, having iterative inputs and in reviewing progress. I also considered, based on my findings, that reflective (Schon, 1983) and organisational learning (Argyris, 1999) approaches combined with strategy, may enhance development effectiveness and governance. A key finding of my research was that "systems ideas" (Williams, 2015) may assist in "understanding interrelationships, engaging with multiple perspectives and reflecting on where boundaries are drawn in terms of those interrelationships and perspectives" (p. 1). Using systems and participatory approaches may provide more holistic and contextual sensitivity for countries, sectors, regions and programmes by including different perspectives (Pawson, 2006). In addition, identifying "invisible mechanisms may contribute crucially" (p. 9) to being more "transformative [and] to sustain social change" (Mertens, 2009, p.18).

### ***Inclusion, partnership and participation***

In the literature review (Chapter 2) I referenced principles such as those outlined by the OECD in *Managing for Development Results* (2008) and in the A/NZ practice context (Chapter 3) for *Managing for Outcomes*. On reflection, after undertaking my review analysis, I considered these principles were positioned more as operational principles rather than as guiding principles or values.

"Guiding values" were presented by Spiller, Barclay-Kerr and Panoho (2016, p. 41). They outlined that "values are positioned as being like the two hulls of the waka (boat)" and were not separate from but form a "values system that creates the mauri ora (wellbeing) in a person, in the group, and in the world" where "values teach people how to adapt and thrive in a changing world" (p. 55). Following this insight, I considered that guiding values of inclusion, partnership and participation could underpin an *evaluative*

*management* approach that is operationalised through practices identified as *strategic evaluation*.

The use of clearly specified guiding values could lead to citizens being more systematically included as partners in steering their development effectiveness and deliberative governance. I suggested their intent as a 'values grouping' reaches beyond concepts such as engagement and collaboration, which were also referred to in my findings albeit with less emphasis. I considered that together these values namely *inclusion, partnership and participation* could be used as 'guiding values' that may enable more authentic relationships and transformational change.

Research participants advocated for increased agency governance and accountability, with more oversight (including an inclusion and equity focus) which could be provided by central agencies. Davies (2007), in his discussion on partnership considered:

that public managers and community activists think in incompatible frames about the role of partnerships and in ways that are not understood by the other party. Non-communication undermines the prospects for an equitable democratic consensus. (p. 779)

In addition, I considered that a mandate for evaluation which incorporated values at a country level could be required, otherwise development effectiveness and governance may remain constrained. These findings mirror increased guidance and mandates for evaluation internationally such as USA, Canada, South Africa Malaysia, Indonesia, Ghana and Uganda.

### ***Integrating theory based evaluative approaches and performance management as strategic evaluation***

Participants supported theory-based evaluative approaches including centralised evaluative designs that enabled more systematic implementation between country and sector levels through to line agencies, regions and programmes. My findings showed that managers and business unit teams at national and provincial levels had an appetite for more theory-based evaluative capacity (including linking logic models, results frameworks and systems approaches) and capability. As an evaluation practitioner, I considered this could mean increased use of strategic results frameworks and alignment at agency, region and programme levels. MFAT's systematic approach of using results frameworks at multiple levels, with indicators identified for aggregating results, and associated guidance (which included evaluative questions and criteria) was an example that participants considered has provided clarity and focus for managers and evaluators.

PNG, with its 2016 mandate for evaluation as a centralised function aligned to the country plan and coordinated by DNPM, also supported this suggestion.

Dahler Larsen (2012) emphasised that “although evaluation is formed and shaped by two large forms of social order (organization and society), the evaluation field has been active, dynamic and almost self-transformative in its responses to challenges and problems”. He outlined that “many evaluative choices take place with reference to values taken for granted in a particular era” (p. 227) and that performance management as a practice was limited by an absence of values. In support of his position and the evidence from participants, I argued that by shifting evaluation to a values-based *evaluative management* approach, the gap between evaluation and management practices can be closed.

### ***Integrating adaptive management approaches and portfolio-based funding***

The fourth of the concept groups was around the use of a country plan by GoPNG and the associated 2016 legislated mandate for evaluation. This indicated recognition that embedding evaluative approaches at central agency level for development and management planning and reporting may be key to achieving more effective development outcomes and systematic data collection.

GoPNG uses a portfolio budgeting approach allocating the budget under sectors against their country plan. Participants considered this may enable more evidence-based and transparent prioritisation and decision-making in the mid to longer term. However, this requires a commitment from all levels of government in ensuring the allocated funds reach the intended targets.

PNG’s mandated cascading planning framework appeared to provide a basis for systematic dialogue on country and regional priorities. As part of each annual budget process, all 22 provinces were required to submit provincial plans which prioritise areas for funding under the development budget. Line agencies, DNPM and PNG Treasury used these plans to assist their portfolio-based budget allocation. However, participants considered some disjunction remained between government policies and funding of service delivery, particularly in health and education.

In A/NZ, the central agency managerial devolution to line agencies appeared to lead to a more fragmented and less inclusive approach for regions and communities. A/NZ research participants considered the inclusive and localised approaches used by the *Whānau Ora* (health and wellbeing) programme was an exceptional initiative that used culturally appropriate and participatory approaches to work effectively at the local level.

However, A/NZ participants suggested that the way funding was approved by A/NZ Treasury lacked transparency and coherence. These findings concurred with a conclusion made by Chimhowu et al., (2019) in recent research on country development planning approaches and processes:

The least convincing area is the way the plans will be financed. It is not clear from our analysis how far this is an issue of low capacity to cost and finance plans, or of a more general lack of access to finance for development. Clearly, this is an area that will require much attention going forward. (p. 87)

Issues such as those raised by A/NZ participants about the transparency of NZ Treasury's annual budget decisions raised a question as to whether A/NZ might consider developing a country plan. Dalziel and Saunders (2014) suggested:

there is a new opportunity for A/NZ to pioneer a further transformation [earlier transformation referred to a welfare state] in how a country enhances the wellbeing of its people, which we refer to as a shift from a 'welfare state' to a 'wellbeing state'. The fundamental difference is where agency is thought to lie in a welfare state; it is accepted that agency lies primarily with central government and the public service; in a wellbeing state, agency is conceived as lying primarily with the country's citizens. (p. 13)

Efforts are now underway in A/NZ to coordinate activities under a wellbeing budget allocation process (Social Investment Agency, 2018; The Treasury, 2019). This new budget direction aligns with A/NZ participants advocating for increased country-level coordination and prioritisation especially in transport planning, agriculture, sustainability and quality of A/NZ water, and the impacts of increased tourism on the environment.

### ***Management and evaluation capability***

My findings signalled that *evaluative management* capability and capacity for both managers and internal evaluators needed to be increased. Participants highlighted that evaluators need to increase their understanding of agency business needs and adapt their evaluative approaches. Participants reported an increased demand for a breadth of timely, evidenced-based results data rather than for evaluators to focus predominantly on in-depth evaluation studies. The findings also suggested that managers required increased knowledge and skills in the use and application of evaluative tools such as results frameworks and intervention logics. In addition, research participants noted that managers needed to be able to develop and adapt these evaluative tools in collaboration with stakeholders.

My analysis highlighted that there was currently no central mandate for evaluation in A/NZ across government agencies, unlike PNG where M&E appeared to be explicitly stepped up at central agency level. Further attention was required within PNG as to how a centralised M&E approach might more effectively cascade to line agencies and regions to enable substantiation of development results. My findings also suggested that with managers acquiring expanded *evaluative management* knowledge and skills, along with an increased use of evaluators, more collaborative and adaptive governance could be undertaken by providing oversight groups within sectors, agencies and regions. These expanded *evaluative management* approaches could be underpinned by ‘guiding values’ of inclusion, partnership and participation. Inclusive and integrated approaches were supported by Innes and Booher (2018) who considered “collaborative governance is emerging in practice in the interstices and across the boundaries of organizations, jurisdictions, and scales of government” (p. 198) where diversity, interaction and methods for selecting actions were emphasised.

### ***Community inclusion***

Finally, in terms of the concept groupings, a key finding of this research was that PNG and A/NZ managers and evaluators (particularly within line agencies) emphasised there was a need to explicitly consider how communities and citizens could be included in more participatory approaches. Jones (2000) in her paper *Partnerships: a common-sense approach to inclusion* suggested that “a connected partnership has potential as a client-centred partnership prioritizing outcomes which focus on the project participants themselves, as opposed to the partners [as] this places the individuals, groups and communities themselves in the ‘subject’ position” (p. 3). This means that rather than line agencies being processed focused, the focus shifts to the partners and citizen requirements, and how the agencies can support improved development results. Participants reinforced this shift may require additional flexibility for processes and contextual sensitivity so inputs can be tailored and adapted.

Jones’ view was supported by Renzio & Lakin (2019) who promoted how concepts such as “equity, sustainability, effectiveness, and inclusion” (p. 19) can be integrated into decisions and budgets.

The focus on efficiency would be complemented by an emphasis on effectiveness in service delivery, to ensure public spending delivers on key results such as addressing basic needs and promoting equality. And the processes and institutions through which budget decisions are taken would be redesigned to become more inclusive, democratic, and participatory, so as to facilitate better-informed dialogue and deliberation. (p. 19)

My findings showed that values of inclusion, partnership and participation were held to be important by most of my participants. These values can be applied (i) vertically from communities, to agencies and sectors and at country level with feedback loops back to the local level and communities such as youth training to employment pathway programmes located in multiple PNG locations; (ii) horizontally within regions across sectors, such as transport planning in A/NZ, which involves local communities, regional and local councils, provincial administrations and the private sector; and (iii) within and across line and central agencies, such as education, where there are differences in educational requirements depending on the diversity of population groups.

### **8.3.3. Institution layer**

My theory-building analysis highlighted three key concept areas within the *institution* layer that may enhance development effectiveness and governance. These were: greater integration of development, management, evaluation and technology functions; strategies that enable an interpersonal organisational culture for feedback, learning and performance, and decision-making; and increased use of mixed-methodologies for data collection, greater use of evidence-based analytical insights, and integrated reporting.

#### ***Integration of development, management, evaluation and technology functions***

First, the research pointed to a need for the integration of development management (strategy, planning, performance management, project management and reporting approaches and processes) with strategic evaluation functions under an integrated 'evaluative management' approach. Such a move may assist in streamlining agency duplication of functions (such as performance management, strategy and reporting) from parallel functional processes such as strategy, policy and performance reporting. Dobell (2003) suggested underpinning these integrated functions with "value orientations" (p. 46) including inclusive and participatory approaches enabled by technology advances (such as digital platforms which can support multiple stakeholders to dialogue and share practice from different locations).

The central question for the moment seems to be whether it is possible that e-governance, based on the ICT revolution, could offer a way out, could re-establish in some sectors sufficient trust in inclusive guiding processes that people would be willing once again to rely on specialized disciplinary or technical expertise (including local and traditional knowledge and other ways of knowing) in implementation of collective decisions pursuing agreed broad value orientations. (p. 46)

This was an area I identified during my research and is highlighted in Chapter Ten of this thesis for further research.

### ***Enabling an interpersonal organisational culture for feedback, learning and performance, and decision-making***

Second, my research pointed to the need for an increased emphasis on embedding learning and performance with feedback loops built-in as part of a system that is integrated into organisational cultures. However, defence mechanisms such as outlined by Argyris (1999) were noted amongst both managers and evaluators in PNG, and A/NZ contexts and these appeared to hinder organisational learning. Argyris suggested using “liberating alternatives [where] organizations and societies that endow human beings with competencies to reverse and undo self-fuelling, anti-learning [and] overprotective processes” (p. vi). This could mean that agencies needed to shift focus from being risk-averse to emphasising learning and focusing on outcomes as a shared culture. This may enable communities of practice where staff and stakeholders contributed as partners for enhancing development effectiveness with governance group oversight supporting practices that are equitable and inclusive.

### ***Mixed-methodology, data and evidence, analytical insights, integrated reporting***

Third, the findings pointed to the value of mixed-methodological (that is, qualitative and quantitative research methods) design. My thesis is that such a multi-dimensional approach to data collection and analysis will be central to an ‘evaluative management’ approach. Research participants considered mixed methodologies make it easier for researchers to incorporate a broader range of different stakeholder perspectives and were more capable of producing “sound, accurate and fair assessments of program impacts” which was seen as “credible evidence” (Donaldson, Christie, & Mark, 2009, p. 46).

My analysis suggested there is a predominant focus on operational processes rather than development results measured by mixed-method quantitative and qualitative data as depicted in the dimensions included in the research impact model (drawn from the OECD (2005a) development and aid effectiveness guidance), accountability document review and by the research participants. Limitations by A/NZ and PNG participants were noted on the analysis and use of evidenced-based development results data. Inequalities were unlikely to be noted or addressed when ethnic and regional population data were not systematically included. This issue was noted particularly in the A/NZ context concerning the needs of Māori, Pacific, and regional population group needs. Participants considered these inequalities were evident within current strategic intent,

agency operations and the uneven achievement of democratic development results. My findings suggested an underpinning use of guiding values such as inclusion, partnership and participation may assist a shift from transactional to more equitable and transformational approaches to improve development effectiveness and governance.

## **8.4 High-level principles**

Further analysis and distillation of the findings led me to consider that guiding values on their own may not be sufficient to ensure that a shift occurred from a focus on operational processes to achieving more inclusive transformative change for people in different situations. I considered that values could guide peoples' practice, yet from my findings there appeared to be other considerations such as different contexts which may require sensitivity and being able to adapt responses to meet changing contexts and peoples' needs. I realised that what was still missing were high level (as opposed to operational) principles in addition to guiding values. One of the initial aims of my research (Chapter 1) was to identify what I described as "higher level" principles to enhance knowledge and practice. I considered that the principles and guidance from the OECD (2006a) and *Managing for Development Results* (2008) were more operational in focus rather than considering, the *wider infrastructure* settings that impact on equity and transformation.

Thinking beyond the findings of my research, I have considered what high-level principles might be relevant to further the conceptual thinking around development planning. Using the findings (both interview and documentary data), I have identified three high-level principles: relationality, contextual sensitivity, and adaptive response, that could overarch the identifiable concept groupings discussed in this chapter. It is the contention of this thesis that these three principles could underpin the ethical, theoretical, and instrumental practices of development – from the architecture and the use of strategic results frameworks to the achievement of development outcomes. Such an overarching conceptual system could provide a series of touchpoints for managers and evaluators to use to understand more clearly the organisational management and governance systems that are engaged in development processes. A discussion of the three high-level principles followed.

### **8.4.1. Relationality**

The first principle I proposed was *relationality*. Under this term, people and their relationships were the central focus. It encompassed an examination of needs, relationships, and communication, but also gave space to examine or consider the politics of inclusion, power, control and devolution. My research highlighted that people in any development context (as both generators and recipients of policy innovation) are



involved across layers of social reality (such as communities, regions, sectors, agencies), in multiple, overlapping processes (including complex and competing ethical claims). People have different perspectives, are from different communities, display human dynamics in agencies and sectors, are involved in a balance between centralisation and devolution of power and control, participate in partnerships and are involved in the sustainability of development and performance. Regardless of a person's social position, they are 'citizens with a stake' and often have strong individual and collective interests in how policies are developed and enacted around them. Those who operate as policy decision-makers in government agencies hold greater power than those who will experience the impacts of policy. A stronger focus on relationality between the whole array of stakeholders could assist agencies to make more deliberative shifts towards considering the needs and demands of stakeholders who are currently marginalised from policy processes.

As Eyban (2008) suggested, what is needed were "more relational processes rather than the things which the processes have generated" (p. 45). This principle would also reinforce Dormer and Gill's (2010) view of the importance of "culture/cognitive – involving shared understandings and logics of actions within institutions which can be downplayed, or omitted if rules and social obligations are operating norms" (p. 1). However, the shared understandings particularly in the A/NZ context, evident in my research, were typically just operating norms as some key stakeholder groups appeared excluded and that there were power and control imbalances between institutions, providers and customers. Therefore, a shift of emphasis onto people, cultures, relationships and perspectives by agency staff may be required (rather than a focus on institution processes). This shift may enhance progress towards transformative development results (social, economic and environmental), that in turn consider sustainability (including wellbeing and resilience) by using more inclusive and relational processes.

#### **8.4.2. Contextual sensitivity**

A second high-level principle, I proposed is that of *contextual sensitivity*. This term implied an awareness and inclusive consideration of cultures and the environments where development planning and management practices were undertaken. While it goes without saying, that people live their lives and interact within specific cultural, geographic, political or other contexts, sensitivity to such contexts requires more than an awareness of this. Sensitivity, in the sense that I proposed requires a *capacity to respond* to the ways in which these contextual differences have an impact on people's abilities to experience wellbeing. Diversity (whether demographic (ethnicity, gender or age and so on), cultural, or geographical) manifests across each cascading layer from country, sector, agency, region, community, programme and project levels which produces

significant challenges for agencies when there was a real expectation that these kinds of diversity are fully attended to. There were multiple layers that appeared to require contextual sensitivity including: ideological and conceptual underpinnings of development and the *wider infrastructure* settings influencing governments; sectors and communities; the architecture and use of strategic results frameworks; and, at the *institution* level, management functions, structures and processes. The word sensitivity was specifically included as it is an evaluative dimension where a judgement could be made. This supports Scriven's (1991) view of evaluation of assessing the merit and worth of interventions or strategies for stakeholders within a context.

#### **8.4.3. Adaptive response**

A third high-level principle put forward from this research was *adaptive response*. This term, as I propose it, implied that the development planning, inputs, implementation and development results might need to be considered within an adaptive innovation systems" approach (Kraemer-Mbula & Wamae, 2010, p. 97). This approach concurred with findings in my research. For example, the Department of Conservation in A/NZ applied adaptive approaches in the way it worked between engagement at the local level through to its institutional learning system and drawing out learnings for innovation and sustainability (A/NZM6).

The theme of resilience reoccurred throughout my research. Participants described resilience as a feature of how people, communities, regions, sectors and countries considered the need to adapt to climate change and catastrophic natural events (such as earthquakes). Resilience was seen as both driven by people's adaptive capacities and as a property of adaptive systems. More research may be needed to tease out the ways in which resilience was core to this concept of *adaptive response* in development contexts.

My theory-building research indicated that systems approaches to adaptation need to be iterative, premised on active feedback loops, and involve people directly in transparent change processes. Establishing strategy, strategic intent, outcome goals and management processes all required responsive adaptation. This *adaptive response* principle could also be highlighted in an *evaluative management* approach. It could draw attention to the need for managers and evaluators to develop capability to understand the way development goals and processes operate with a complex system. Capability with such an approach could enable timelier portfolio budget allocation, better-tailored services and interventions that are responsive.

## 8.5 Chapter summary

The identifiable concepts drawn from the research included both some that were expected and others that seemed emergent. Overall, the direction indicated by the findings led to the development of a theory-building idea for both repositioning evaluation as a core component of strategic organisational function and the increased integration of evaluation as a corporate management function under a paradigm shift towards *evaluative management*. My research clearly indicated that good practice concepts do not or cannot operate systematically and effectively in the original hierarchy of layers as depicted in my research conceptual frame (see Figure 5).

The discussion in this chapter focused on the extent to which the substantive and emergent concept areas identified through the analysis provided an extended knowledge base for understanding good practice concepts in the architecture and use of strategic results frameworks, associated management systems, and governance. This suggested that increased attention needed to be paid to these emergent concept areas and proposing a new conceptual model was one way to achieve this.

In summary, my research indicated: (i) there were emergent concepts that can be drawn from analysis that looked across development, management and evaluation knowledge fields, and (ii) that the nature and positioning of these concepts, particularly *strategic evaluation*, might be usefully reconfigured under a proposed paradigm shift provisionally identified as *evaluative management* which, in turn, may (iii) enable integration and extension of interdisciplinary knowledge and good practice concepts in the context of country, sector and programme results frameworks and associated management and governance systems, and, finally, (iv) included the identification of guiding values and high-level principles to underpin such an *evaluative management* approach.

From here, I refer to this new approach as *Values-Based Evaluative Management*. This approach included the three identified high-level principles of *relationality*, *contextual sensitivity* and *adaptive response* underpinned by the three guiding values of *inclusion*, *partnership* and *participation*.

In Chapter 9, a *Values-Based Evaluative Management* approach was discussed in more detail while Chapter 10 concluded the thesis argument by rehearsing the implications for the use of an *evaluative management* approach and associated good practice concepts, and identified areas for further research.

## Chapter 9: Values-Based Evaluative Management

This chapter addressed research question four: How do the emerging good practice principles in country, sector, agency and programme systems work to enhance development effectiveness and governance for development (national governments and partners), management, and evaluation? (RQ4) This chapter presented a possible way of modelling what needs to be considered in a new *evaluative management* approach (a ‘*Values-Based Evaluative Management System*’) for country-level development. While I discussed the proposed model at the country system level, it should be feasible drawing on a set of good practice principles, values and concepts to apply the model at different levels (such as sector, region, agency, provider, community, programme or project).

One aim of my research was to expand what is known about how evaluative perspectives and practices may be used to enhance development effectiveness and governance. My initial idea was that there needed to be a better way to connect evaluative approaches in management, and then align with the new SDGs in national development settings. The second aim of my research was to enhance development effectiveness and governance by identifying a conceptual good practice framework underpinned by high-level principles. The identified good practice concepts, discussed in Chapter Eight, were now presented within a model that described a *Values-Based Evaluative Management System*, the purpose of which was to support a focus on more equitable and sustainable development results.

The OECD (2009) development effectiveness guidance for using results frameworks was for all development partners (national government, civil society, private sector and donors) to align to a country development narrative through country, sector and region results frameworks. Over the past five years, there was an increasing global focus on sustainable development, which widened development goals to systematically include economic, social, and environmental dimensions. Schwandt et al. (2016) stated that “evaluation must go beyond measurement, to consider whether progress is equitable, relevant and sustainable. Such evidence will help demonstrate public sector accountability and accelerate change by focusing attention on enhancing learning and innovation” (p. 1).

Approaches such as Chen’s theory-based evaluation (1990), Funnell & Rogers’ programme theory (2011); Fetterman’s empowerment evaluation (Fetterman & Wanderman, 2005), *Enhancing Mātauranga Māori and Global Indigenous Knowledge*, NZQA, 2014); Mertens’s (2009) transformative research and evaluation, realist evaluation (Pawson, 2006, 2008, 2013) and Williams (2015) systems were examples of authors contributing to my thinking for an *evaluative management* approach. I considered

an *evaluative management* approach could contribute to supporting 'new' national development planning where Chimhowu et al. (2019) acknowledged "this emergence of 'new' national planning must be seen as an opportunity for the global community to continue finding ways of enhancing the achievement of the SDGs" (p. 87).

My findings pointed to a potential for more efficient and effective service delivery being facilitated through a renewed strategic focus that appeared to be emerging at the interface of development, management and evaluation. Since 2015 (Schwandt et al., 2016), a range of new strategic-level discourses began to emerge in government spaces including how to better enable citizens and beneficiaries to participate in social, economic and environmental development planning, ideas of wellbeing and environmental sustainability, and capacity for resilience in the face of mounting mental health (Dalziel & Saunders, 2014) and environmental challenges. In both A/NZ and PNG, a shift in strategic intent was driven by a range of factors. First, there was an increasing awareness of inequalities in service access and delivery at the regional scale. It was also informed by inclusion and equity issues highlighted by increasing diversity and significant population subgroup inequities. My research reinforced that more explicit strategic intent for different population groups and regions to address these issues was required which involved stakeholders and measurement of progress.

Finally, an increased use of new information technologies appeared to provide potential transformational opportunities to integrate, reconfigure and streamline management functions, including a more explicit role and need for a form of strategic evaluation, within government agencies. Taken together, more integrated and innovative evaluative approaches that can be adapted by stakeholders to different contexts appeared to be needed by government agencies, regional and community groups to ensure inequalities were addressed and transformative results sustained.

Countries like A/NZ which does not have country-level "development plans" may be unable to formulate country-level strategic direction as planning is devolved to core agencies. This also means there is little capacity to develop inclusive dialogue and adaptive processes at the country level. Devolution is, however, one strategy for mitigating the disruption of short political terms. My research suggested that countries such as PNG were able to maintain some implementation continuity through the alternate strategy of a country plan which may also serve to mitigate negative outcomes from short political terms as sustained change is incrementally achieved. Thus, it appeared countries without a national development plan and associated good practice concepts within each of the layers depicted in Figure 14 and Table 15 may find development effectiveness and governance constrained. However, in either context, mid to long term

implementation horizons with a more explicit role for strategic evaluation appears required to measure progress towards more equitable and sustainable development.

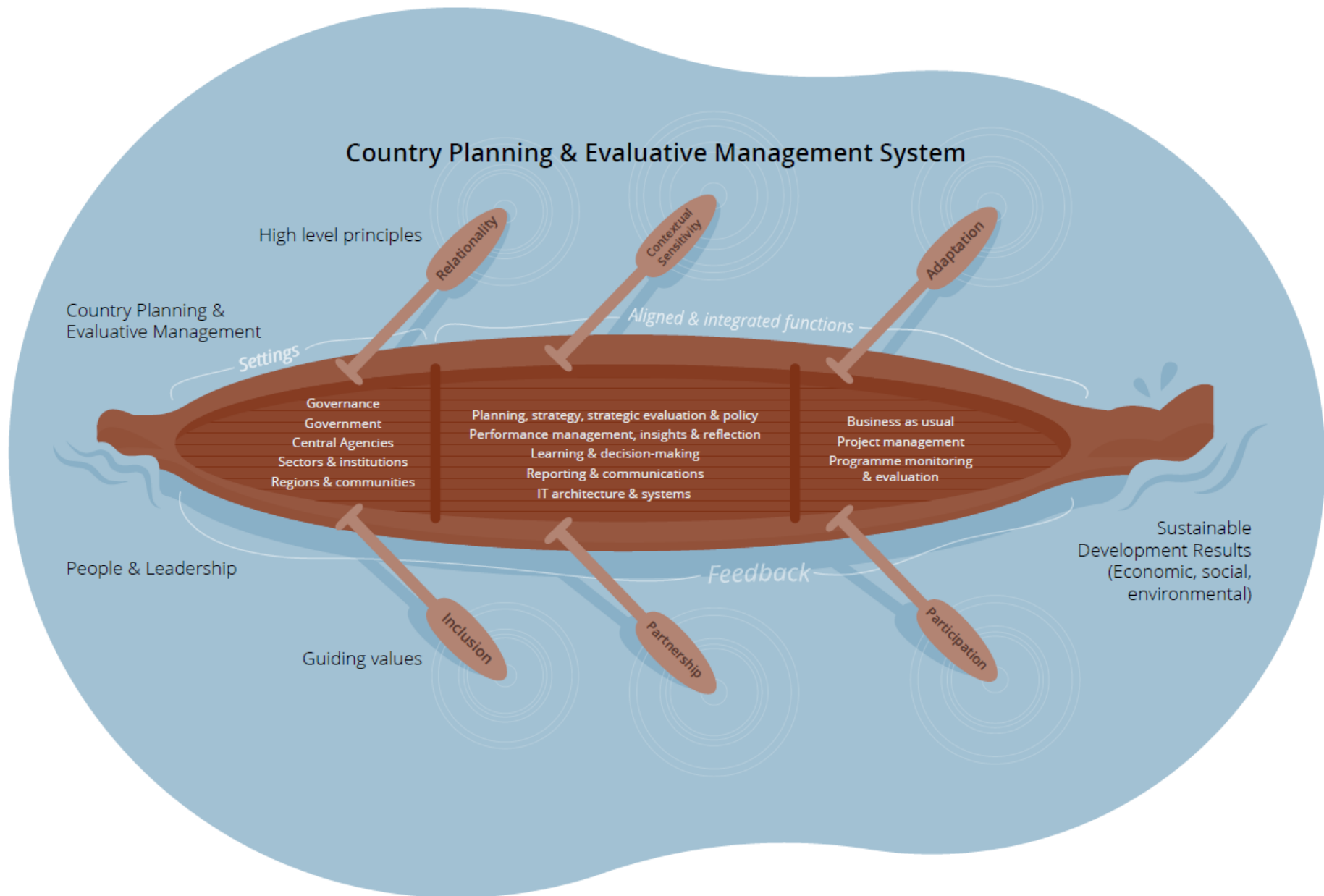
My research, coupled with my own experience as an evaluation practitioner in the development field, suggested that a shift in worldview (Vidal, 2008) where a view that considered values and approaches may now be required. In the next section, I proposed one way such a new worldview might be conceptualised and described it as an *evaluative management approach*.

## 9.1 A new model

The new model developed from this theory-building research is displayed in Figure 13 and is described as a '*Values-Based Evaluative Management System*' (VB-EMS). It displayed an integrated and adaptive approach to enhance inclusion, development effectiveness, governance and sustainability. The model depicted an active system of *evaluative management* as a knowledge discourse and practice approach embedded in good practice concepts (Table 15).

During the analysis and reflecting on my distillation of the key good practice concepts and the role of guiding values, I began to explore boats as a potential metaphor for how to describe evaluation as a tool for steering strategic direction. Thinking about this in an A/NZ or PNG context led to thoughts about *waka* (canoe) or *kanu* (Māori or Tok Pisin respectively) to visualise this idea of a boat. Spiller et al. (2016) discussed values in relation to a waka where "values are positioned as being like the two hulls of the waka" (p. 55). The importance of wayfaring responsive leadership was emphasised. Barclay-Kerr (2006) outlined "how Aotearoa (New Zealand) waka became predominantly single-hulled vessels... determined by the type and quantity of native trees" (para. 2). A Māori metaphor commonly used in A/NZ is "He waka eke noa – we are all in this (waka) together (Kupu Māori, 2019). Another metaphor also used emphasises that real leadership requires to get everyone to work together "kia kotahi te hoe o te waka – to literally paddle as one". King (2007) highlighted how the metaphor of a canoe incorporates a Māori worldview, linking language and, and refers to getting on board for a journey (pp. 155-157).

These metaphors encapsulated key elements of my thinking and provided an incentive for me to consider a canoe analogy further. I decided to depict one hull in my model with values and high-level principles as key constructs depicted by the six paddles which guide the journey. As highlighted in my research, country development is enabled by people and leadership (guided by values and principles) contributing within different settings (vertically and horizontally) together undertaking aligned and integrated functions based on the associated good practice concepts listed in Table 15.



**Figure 13. Country Planning and Evaluative Management System Model.**

**Table 15. Values-Based Evaluative Management: guiding values, settings, high-level principles and associated good practice concepts.**

Country Planning and Evaluative Management System		
Good practice concepts		
<b>People and leadership</b>	<b>High-level principles relationality, contextual sensitivity, and adaptive response</b>	<b>Feedback</b>
	<b>Wider infrastructure context</b>  <b>Development:</b> <ol style="list-style-type: none"> <li>1. Governance systems</li> <li>2. Leadership (horizontal &amp; vertical)</li> <li>3. Development planning &amp; accountability</li> <li>4. Centralisation, devolution, and regionalisation</li> <li>5. Power, control &amp; ownership</li> <li>6. Sustainability, wellbeing &amp; resilience</li> </ol>	
	<b>Institution context</b>  <b>Management:</b> <ol style="list-style-type: none"> <li>13. Managerial evaluative capability</li> <li>14. Citizen, beneficiary, and customer voice</li> <li>15. Interpersonal organisational culture</li> <li>16. Strategy and business models</li> <li>17. Aligned and integrated management functions (including planning, strategy, organisation and programme evaluation, policy, business as usual, project management, performance and financial management and reporting) and processes with an outcomes and impacts focus</li> <li>18. Feedback, reflection, and learning</li> <li>19. Decision-making</li> <li>20. Adaptive management</li> </ol>	
	<b>Strategic evaluation:</b> <ol style="list-style-type: none"> <li>7. Needs and capability assessment</li> <li>8. Strategic intent and portfolio-based funding</li> <li>9. Theory-based evaluative approach (including alignment of country and sector strategy, planning, policy, performance management, reporting, communications, and learning)</li> <li>10. Iterative mixed methodology</li> <li>11. Data insights &amp; feedback</li> <li>12. Evidence-based reporting and communications</li> </ol>	
<b>Guiding values – inclusion, partnership and participation</b>		

A key insight of VB-EMS was that what previously was identified in my analytical adapted research analytical frame (refer Figure 5) under the *ideas, interventions and frameworks* layer appeared to need repositioning in the new model under evaluative management and, in particular, as *strategic evaluation* needs to be identified separately. The idea behind the repositioning – in effect bringing the *ideas* layer higher in the nested hierarchy of contextual reality layers – was that this positioning foregrounded values, principles and associated good practice concepts in the management rather than the implementation sphere. Such framing has the potential to conduce management to



embed a values and principle-based approach to enhancing development effectiveness and governance. Promoting values of inclusive, authentic partnerships and iterative participation, I argued, could support more equitable and democratic processes for citizens, and in turn, enable more inclusive sustainable development and governance.

Positioning *evaluative management* as an idea that needed a place within management discourse was also something that I argued. Uptake of the idea would help address governance and management (including performance accountability oversight gaps at a national level, which were knowledge gaps highlighted in the literature review and the qualitative research findings). Country-level governance, development and management can then be articulated through a 'worldview' of 'guiding values' of '*inclusion, partnership and participation*' promoting deliberative and more democratic governance and sustainable development.

My research contributes to providing a country development narrative which highlighted values, principles and good practice concepts, that in my view requires consideration. This approach may enable more sustainable development results that can evolve iteratively, as countries navigate their journeys over time through changing conditions which may involve altering course whilst supporting the wellbeing and resilience of its people. In the bow of the waka (Figure 13), country planning and *evaluative management* requires people and inclusive leadership from the different settings to set the country direction and guide the way based on values and principles. A deliberative governance oversight of a country and settings is used with layers of government (including ministers elected by the citizens), central agencies, sectors and institutions (agencies, regions and communities) having a relational role in working together. This can be undertaken in partnership, with contextual sensitivity through participation and adaptive responses working towards more equitable and transformative sustainable development results.

The aligned and integrated functions involved two groups working together as depicted in the model (Figure 13) including (1) planning, strategy, strategic evaluation and policy, performance management, insights; and reflection, learning and decision-making, and reporting and communications; and (2) operational business as usual, project management, programme monitoring and evaluation; and IT architecture and systems with feedback loops. The integration and alignment within these two groupings may require further research. Streamlining functions within these functional areas could further enhance practice.

Packer (2011) suggested that “the linkages between formal knowledge and embodied, social know-how” (p. 13) were more critical than has often been considered. The good practice concepts distilled in my research were central to the idea of *evaluative management* as indicated in Table 15. They existed across the conceptual fields Pawson (2006) described of *infrastructure* and *institution* but also in the context that my research identified of *strategic evaluation*.

## 9.2 Good practice concepts

The following table (Table 16) provides a brief explanation drawn from my research for each good practice concept group under a *Country Planning and Evaluative Management System* approach. These concepts were grouped (as depicted in Table 15) under development, strategic evaluation and institution management knowledge areas.

**Table 16. Good practice concept explanation under a Country Planning and Evaluative Management System approach**

Concept group	Explanation of concept group that may enhance development effectiveness (drawn from this research and evaluation practitioner ‘lived’ experiences)
<b>Development concepts</b>	
Governance systems	Oversight, by a representative group focused on inclusion, equity, scope, and management activities within a specified system boundary of sub-system components and relationships.
Leadership (horizontal & vertical)	This Involves leadership by horizontal groups from multiple functional areas (such as planning, strategy, strategic evaluation and policy, performance management, insights and reflection, learning and decision-making, and reporting and communications) and vertical leadership groups which includes managers from different contextual layers (such as central agency, institution (line agencies), region, community and customers, beneficiaries, and citizens). These groups particularly the vertical leadership groups (human agency ‘situated in the bow of the waka’) may enable more adaptive responses to changing contexts.
Development planning & accountability	This involves undertaking collaborative national planning and assessing progress (for inclusion, equity and sustainability) of a country, sector, region or community using both quantitative and qualitative data to provide insights and learning.
Centralisation, devolution and regionalisation	This includes use of a centralised government role with representative regional perspectives and functions. This includes the sharing of management responsibilities between national and regional stakeholders that represent their local contexts and changing needs.
Power, control & ownership	A values-based worldview that includes, empowers and enables equity at country, sector, region, and community levels.

Concept group	Explanation of concept group that may enhance development effectiveness (drawn from this research and evaluation practitioner 'lived' experiences)
Sustainability, wellbeing & resilience	Sustainable development encompasses social, economic, and environmental dimensions, including wellbeing and resilience for people and the environment.
<b>Strategic evaluation concepts</b>	
Needs and capability assessment	This involves a collaborative assessment of local contexts and different peoples' requirements, and supports the skills, knowledge, capability, and capacity required to participate as partners.
Strategic intent and portfolio-based funding	This includes country and sector goals and strategies with associated theories of change linked to overarching policies, with funding allocated for implementation across multiple sectors, line agencies and programmes
Theory-based evaluative approach	This incorporates development and use of a country and sector approaches including alignment of country and sector strategy, planning, policy, performance management, reporting, communications, and learning functions. The use of a strategic framework based on country and sector theories of change is key as a dialogue and collaborative tool that can then be aligned policies and data collection, effectiveness assessments undertaken at country, sector and regional levels with associated reporting and communication for learning and performance.
Iterative mixed methodology	This includes both quantitative and qualitative data collection providing the 'what' and 'why' using shared digital platforms for 'real time data collection where feasible.
Data insights & feedback	Data is segmented for ethnicity, gender, regions and adaptive needs, and analysed and used to provide into inputs into strategic and progress assessments.
Evidence-based reporting and communications	Data from multiple sources are included in assessing the progress of theories of change and policies, which are used in reporting and communications.
<b>Institution management concepts</b>	
Managerial evaluative capability	This includes building managers' skills and knowledge to develop and use strategic frameworks encompassing theories of change, measure results, use both qualitative and quantitative data, and assess progress that is equitable and sustainable.
Citizen, beneficiary, and customer voice	Stakeholder and people's views and perspectives are included throughout the design, planning, implementation, analysis, and reporting phases.

Concept group	Explanation of concept group that may enhance development effectiveness (drawn from this research and evaluation practitioner 'lived' experiences)
An Interpersonal organisational culture	Government agencies ensure people communicate, share and respect different views and inputs throughout the integrated approaches, processes, and learning.
Strategy and business models	Government agencies clarify their purpose, goal, and operations for working together and achieving equitable and sustainable development results for citizens while preserving the natural environment.
Aligned and integrated management functions	Management functions (including planning, strategy, strategic and programme evaluation, policy, business as usual, project management, performance and financial management and reporting) and processes with an outcome and impact focus are process mapped and streamlined through values-based (inclusive, partnership and participatory) outcome-focused and integrated approaches.
Feedback, reflection, and learning	Values-based analytical processes are used, and the evidence-based learning is considered and shared.
Evidenced-based decision-making	Decisions using data and learning are undertaken, cognisant of equity and sustainability.
Adaptive management	Inputs and processes are iteratively evolved cognisant of contextual changes and peoples' needs to enhance management practice and results.

The above concepts are intended to enhance management practice and achieve more equitable sustainable development results and extend knowledge of the potential of evaluation in the development field. I argued, based on research participants' responses, that there was a gap in current practice approaches, which therefore provides an opportunity for something like a *Values-Based Evaluative Management System* approach. I argued that grounding a development approach in high-level principles that are values-informed has the potential to lift practice from being operational to transformative.

### 9.3 Chapter summary

This chapter presented a model of an integrated country planning and evaluative management system. The model shows there is an interdependency between development, management and evaluation concepts that have not been clearly articulated to date. The suggestion of asking managers, evaluators, agency teams, service providers and communities to consider all of the elements is a tall order, and the idea of the visualisation is to help compress these ideas into a diagram that can be used as a starting point for discussion about more innovative evaluative approaches to

development challenges. Ideally, good practice, high-level principles of relationality, contextual sensitivity and adaptive responses could become everyday knowledge and practice for all people involved in country-level development. In Chapter Ten, conclusions from this research and implications for the theoretical discourse and practice were highlighted, and areas for future research identified.

## Chapter 10: Conclusions and implications

With growing awareness of climate change impact and repercussions from adverse environmental and political events, stakeholders are calling for increased inclusion, equity and downwards accountability to citizens with increased horizontal accountability at different management levels. Coupled with this is the desire, expanding globally, to modify human practices towards more inclusive, sustainable development. My thesis was that current evaluative approaches to the management and governance of development, particularly in the context of the SDGs, were insufficient to ensure robust, inclusive, and effective results. The overarching research question that I have focused on to explore this argument was: *What are 'good practice' evaluative principles and concepts that may enhance country, sector and agency management and governance knowledge and practice in different contexts?* In this chapter, I present conclusions, discuss the implications for country development, management and evaluation, and identify potential implications from my research.

### 10.1 Conclusions

My research was situated within a development and strategic management paradigm for evaluation. This theory-building research was underpinned by a critical realism theoretical perspective. I used an emergent conceptual research typology which was initially adapted from Pawson's 2006 frame and then further developed during the scoping phase of this theory-building research. I incorporated interpersonal relations and individual capacities into each of the three contextual layers – *wider infrastructure*; *institution*; and *ideas, interventions and frameworks* – as I found in my scoping phase these two dimensions to be more widely relevant than Pawson's (2006, 2008) suggestion of them as part of four additional contextual factors (Pawson 2004, p. 7). As a result, I have presented them as situated within the layers of *infrastructure*, *institution*, and *ideas, interventions and frameworks*. This innovation allowed me, in turn, to construct a narrative about the importance of underlying values of inclusion, partnership and participation, and the centrality of relationality, contextual sensitivity and adaptive responses as overarching principles.

In my analysis (see Chapter 4: Methodology), I used the critical realist strata of "experiences, events and mechanisms" (Wuisman, 2005, p. 384) I utilised theory-building to examine the differences between the two case study findings and identify associated good practice management concepts that could be applied in different country contexts.

I observed, from my secondary research analysis that the aspects of my thematic findings appeared to be the identifiable good practice concepts. I considered that the application of the good practice evaluative concepts by people and leaders within the three contextual layers of *infrastructure*, *institution* and *ideas* may be central for more equitable and transformative sustainable development results (social, economic and environmental).

Applying the contextual layers (in their modified form) in a country development context proved a useful methodological approach, as the research helped to deepen my understanding of country-level development systems in PNG and A/NZ. Including both managers and evaluators in my qualitative research extended my understanding of the findings. The rating matrices of good practice concepts derived from the accountability reporting documents and self-rating by A/NZ managers and evaluators was useful because it helped me to identify recognised and emergent good practice concepts from my research. The knowledge and practice gaps became more apparent during my analyses and laid the groundwork for identifying which good practice concepts required increased attention and inclusion in development, management and evaluation knowledge fields as part of my theory-building research approach.

Based on my previous academic studies and practical experience, this theory-building research generated five elements that were surprising to me, and these led to the iterative conclusions from this research. First, I realised that good practice concepts could be identified from within each of the three layers (that is, *wider infrastructure*; *institution*; and *ideas, interventions and frameworks*). My research findings suggested that these concepts were wider ranging than previously documented in the literature and practice guidance.

Second, I considered that the good practice concepts identified in the *ideas* layer could not operate if evaluation remained at an operational level within institutions. Hence, I argued that evaluative approaches needed to become more strategic and designed at a central agency level. This view was also in line with trends internationally, which included mandating a role for evaluation at the country level.

Third, my research analysis showed that PNG, by taking a country approach to development planning, and then with using evaluation at a more strategic level, assisted with setting a country direction and incorporating emerging themes such as sustainable development into this process. My research in A/NZ, however, suggested that a devolved managerial approach to line agencies appeared to create a lack of clarity for managers and evaluators on country and sector goals. In addition, implementation constraints were

evident in relation to underpinning themes such as increasing living standards and wellbeing. Accountability and performance gaps were highlighted by research participants in both A/NZ and PNG.

Fourth, reflection on my findings suggested that a shift to having guiding values underpin an evaluative approach may be required to enhance inclusion, partnership and participation. I considered that these values may need to be made explicit within an evaluative approach to ensure their application to drive more equitable and democratic practices. In addition, I also considered that high-level principles such as relationality, contextual sensitivity and adaptive responses were required to provide an overarching framework for an effective evaluative approach.

Fifth, I then considered that a centralised *evaluative management* approach with guiding values, high-level principles and associated good practice concepts together may be sufficient alongside country development planning for enhancing development effectiveness and governance. The multi-element approach evolved iteratively during my analysis. I was surprised by the combination of these elements and considered that it may be necessary to extend knowledge and transform practice towards more equitable and sustainable development based on a more holistic understanding of development management and governance.

An outcome of this research was a transferable integrated and adaptive model of *Values-Based Evaluative Management* positioned in the management knowledge base that was focused on integrating people and resources, and further enabled by integrated technology.

I argued that a 'world view' underpinned by guiding values such as inclusion, partnership and participation appeared needed as we live in a world that has different social realities and perspectives for each of us. My research has led me to reconsider the neoliberal free-market approach to development and to note the ways in which it has led to inequalities and global environmental concerns. I considered that by sharing guiding values and adopting a values-based approach, more sustainable development and democratic governance might be enabled. The contribution evaluation can make to this resides in its capacity to enable more inclusive, equitable and transformative approaches with associated good practice *evaluative management* concepts underpinned by explicit values and principles.

The research identified three high-level principles – relationality, contextual sensitivity, and adaptive response – and my claim is that these principles were necessary to underpin more robust development systems at country level. It is the contention of this



thesis that these three principles need to be brought into clearer focus because they underpin the ethical, theoretical, and instrumental practices of development and management. While they were identified within the *wider infrastructure* setting, they can be seen to be influential across development and management at country, sector, institution, region, community, programme and project levels.

My research also suggests that combining concepts such as centralisation and devolution by increased overarching centralised development management planning integrated with theory-based evaluation systems at national levels and within this at subnational levels, power and management devolution to regions and agencies can be increased to enabling more self-directed focus and power. This includes measurement of development results using mixed-methodology to provide more insights into citizen perspectives, an integration of management processes enabled by theory-based evaluation strategic frameworks and extended management evaluative capacity to apply the good practice concepts highlighted in Table 16. An increased role for central government in design, evaluative management (encompassing performance) and accountability guidance, with governance oversight at both central and subnational levels, are also required.

I suggest that, without an understanding of the role and potential of evaluative thinking at strategic levels in government management, enacted through such principles and values and well communicated and widely understood, the capacity for effective interaction and communication between different levels of government is severely constrained. Thus, the capacity to address equity, governance, accountability and sustainable development is also constrained.

#### **10.1.1. Contribution of this thesis**

An aim of this theory-building research was to contribute to an evidence base and practice in an interdisciplinary interface between development, governance, management and evaluation. My research challenges basic assumptions of development, governance and management and evaluation as it potentially repositions development, management and discourse for national development of countries, sectors, regions and communities. I consider an underpinning of guiding values and high-level principles of an integrated and adaptive *evaluative management* approach and country system model can enhance development effectiveness and governance. Guiding values of inclusion, partnership and participation together with relationality, contextual sensitivity and adaptive responses are put forward as ways to support more deliberative and democratic governance, development, management, and evaluation practice, which may, in turn, enable more inclusive and sustainable development and

improve lives for citizens. I consider relationships are critical processes for inclusion and accountability between communities, institutions, and governance to enable sustainable development with institution and management systems adapting to more complex and changing contexts.

*The Values-Based Evaluative Management* approach was presented and discussed with managers in both developed and developing country contexts and with evaluators at conferences and in workshops. Currently, evaluation knowledge and skills are regarded by many evaluation practitioners and associations as a specialist area. As an academically qualified and practising manager and evaluator, I consider that these guiding values and high-level principles of *Values-Based Evaluative Management* could be incorporated into management and development theory discourse and form the basis for good practice in development contexts.

Alongside strategic planning, performance management would be strengthened by being repositioned and integrated with strategic evaluation within an *evaluative management* approach. An integrated approach is promoted due to the limitations in approaches and knowledge gaps in the planning, policy, implementation, evaluative and reporting cycles and duplication across different functional areas. Incorporating, relational approaches with enabling technology, and responsive and adaptive management practices may expand knowledge and practice and extend capacity and capability within institutions – central, line and regionally. I consider, based on my 'lived' experiences as an evaluation practitioner (working externally in A/NZ, PNG and other countries, and more recently as a principal evaluator within a government agency for the A/NZ public sector), these values and high-level principles underpin an integrated and adaptive *evaluative management* approach that can be applied at community, institution and central agency levels. I endeavoured to draw together, reconfigure and reposition evaluation through a 'new' integrated management and evaluation paradigm of *evaluative management* as a knowledge and practice discourse that could enhance development effectiveness and governance, and in turn contribute towards more sustainable development.

#### **10.1.2. Limits of the research**

This mixed-method research included document review, a literature review, qualitative interviews (n=48) with managers and evaluators, and a self-rating matrix on practice concepts completed by A/NZ research participants. The research was completed in two case study contexts (PNG and A/NZ). The number of case studies initially proposed was five, including Australia, Samoa and Laos. However, given the complexity of findings that emerged in the scoping phase, the number of case studies

was reduced to two to ensure the fieldwork could be completed. The case studies included a developed country (A/NZ) and a developing country (PNG) with different approaches used for country development. My findings appeared to align with recent international literature such as Chimhowu et al. (2019) on 'new' national planning, incorporating system approaches for evaluation (Williams, 2015) and inclusion of 'coherence' in the recently updated evaluation criteria by the OECD (2019).

A consideration I reflected on during the analysis of my findings and discussed with colleagues included: Does an *evaluative management* worldview support a Māori worldview? Are the values and high-level principles appropriate to consider that *evaluative management* could be used with different cultures and ethnic contexts? As a Pākehā (European) New Zealander, I did not want to presume applicability. When presenting my research findings and an idea of *evaluative management* as a potential 'new' integrated management approach or paradigm at an evaluation conference in 2018, a senior Māori evaluator involved in iwi governance said he considered *evaluative management* could be a relevant and useful approach with its emphasis on values and high-level principles. This view was also supported by a Samoan colleague who acknowledged relevancy with an underpinning of values and principles such as inclusion, relationality, contextual sensitivity and adaptive responses. However, this was an area that needed further exploration.

Other methodology that could have been employed for this study included using an interpretative methodology for this research. However, I decided that although useful from an anthropological approach, it would not provide me with a critical lens to examine what is working for different population groups and levels in government. My aim from this theory-building research was to draw out potential principles and knowledge and practice concepts for managers and evaluators that may be relevant in different development contexts to extend knowledge and enhance practice.

Since this research was completed in 2016, A/NZ government agencies appear more aware of inclusion in planning and service delivery for regions, gender and youth. However, by late 2019, there was still limited guidance and no mandate for evaluation or on systematic approaches for inclusion of key population groups such as Māori and Pacific peoples, disability sector, and provincial needs and priorities.

During 2018 and early 2019 as this research was written up, A/NZ government announced their intentions to shift more service delivery coordination such as education to regional hubs. This supports the findings that research participants considered there was structural misalignment (particularly evident in the transport sector) in A/NZ and an

absence of a coordinated country approach with regional inputs was impacting on achieving and sustaining social, economic and environmental outcomes (such as increased children living in poverty, decreasing water quality, environmental impacts from tourism). These outcomes were compounded natural events such as earthquakes and global warming, and most recently from the mass shooting in Christchurch raising questions on inclusiveness and valuing of diversity in A/NZ.

## 10.2 Implications

From the beginning, my research intended to contribute to development planning and evaluation in different contexts, with a particular focus on PNG and A/NZ and investigate whether this research could contribute to a broader knowledge base. I considered implications from my research under five areas.

First, my research indicated a need for strategic evaluation to be shifted within national development to a central agency role and function with a mandate which already was being undertaken in countries such as Canada, South Africa, Ghana, the UK and the USA. I considered that systematic inclusion, authentic partnerships and participation of citizens in their country's development required an evaluation mandate with additional performance and accountability oversight at central and sub-national levels as part of more deliberate and democratic governance. Mandating *strategic evaluation* as a centralised function and using an *evaluative management* approach premised on values and high-level principles aligned with the global 2030 SDG Agenda assist countries in working towards more equitable and sustainable development results. I argued that without this mandate and oversight, inclusive sustainable development might be constrained, and impact on people and our environment for future generations.

Second, based on my research, I considered that management practices embedded within neoliberal economic settings might have led to inequalities and exclusion for key population groups, regions and communities which required addressing through explicit inclusion, changes to 'rules', potential legislation, and increased guidance by governments for national development. The *Values-Based Evaluative Management* approach developed above could support shifts to more sustainable development underpinning development planning (such as in PNG) and 'wellbeing economics' (Dalziel & Saunders, 2014). This includes such approaches that A/NZ is underway with (such as the *Wellbeing Budget*, The Treasury, 2019). This shift of emphasis is not only for people but also applies to preserving and restoring our environment and supports greater participation of citizens as 'partners' in their community, regional, sector and country development. This may require shifts in balance of power and control with increased recognition and empowerment of local knowledge, cultures and practices.

Third, my research indicated that countries, sectors and regions require more inclusive development planning processes, which include increasing their collaborative dialogue and using more inclusive and evidence-based decision-making, centred around a form of *Values-Based Evaluative Management*. This could involve developing plans considering multiple perspectives, using inclusive approaches in goal setting and activity design, and measuring progress with active feedback loops and reflection, and integrating learning. My findings highlighted that it is important for countries to be cognisant of the balance of power and control that result from centralisation and devolution processes. Therefore, 'decentralisation and regional effectiveness' could be added as dimensions within development effectiveness, with a governance oversight potentially undertaken by central agencies at country, sector and regional levels. This approach could provide regional development analyses and more deliberate and democratic governance within country development.

Fourth, my research highlighted that increased country development intent and governance is required. Central agencies could first consider providing country direction through partnering with governance teams from sectors and regions to collaboratively decide how a country can own its development in more deliberative and democratic ways. Moreover, I argue that without strategic leadership, direction and governance, constraints in inclusion and participation of citizens may continue to occur.

Fifth, I considered, based on this research and from my evaluation practitioner 'lived' experiences, an increase in specific managerial *evaluative management* knowledge and capability based on the good practice concepts highlighted in Chapter 9 (Table 16) was needed. For specialist evaluators, an increase in strategy and business knowledge and practice appeared required so they can co-design and undertake more specialist roles at a strategic level. For evaluation associations, this could include an *evaluative management* stream to support practitioners. Universities could teach *evaluative management* integrated into their strategic planning and performance management discourse within their management and development courses. A knowledge and practice stream of *evaluative management* offers an opportunity for the evaluation profession to contribute to this knowledge and practice gap.

Finally, other areas that require further consideration arising from this research and other international trends include national development and the contribution that *evaluative management* could offer in governance oversight, co-designing of plans, measuring progress and reporting. Another area for further consideration and research was the potential integration of management functions enabled by technology (such as strategy, planning, performance management and reporting) and alignment of policy,

project management and operational monitoring and evaluation, which may assist in streamlining institutional functions and processes, thus reducing duplication, and enhancing adaptive approaches.

Yet, my findings also indicate that there are embedded paradigms of management practices entrenched within 'top-down' strategic planning and 'tunnel vision' quantitative performance management, with limited attention to relational and adaptive approaches and processes. I consider guiding values of inclusion, partnership and participation, and high-level principles of relationality, contextual sensitivity and adaptive responses underpinning an *evaluative management* approach with integrated good practice concepts are needed to contribute to a transformational shift towards more inclusive sustainable development and governance.

The *Values-Based Evaluative Management* paradigm proposed in this **theory-building** research requires a commitment to a country development mandate, expanded guidance, and learning and accountability oversight to support achieving the articulated goal of more inclusive sustainable development. I consider that both management and evaluation associations could have major roles to play in the coordination and the embedding of *evaluative management* knowledge and practice across institutions and communities, and within management theory discourse and practice.

Given the diverse knowledge concepts highlighted from my research in *evaluative management*, a focus on extending the knowledge and skills of managers and communities to enable participation in more authentic and equitable partnerships is required. I consider, based on this research and my 'lived' experiences as an evaluation practitioner, shared values and high-level principles with good practice concepts can promote and enhance knowledge sharing and lead to more equitable, adaptive and responsive decision-making. This, in turn could enhance development impacts, development effectiveness and governance by governments, organisations and communities, and enable more empowered and engaged citizens in development, and enhance peoples' lives.

**Sub-note:**

During late 2019 and 2020, further public sector reform efforts were underway, led by the A/NZ State Services Commission, on mandating inclusion for Māori in all aspects of the public service. In August 2020, the 1998 State Sector Act was repealed, and the Public Sector Act 2020 (SSC, 2020) was passed.

The Act provides a modern legislative framework that enables a more adaptive, agile and collaborative public service and includes stronger recognition of the role of the public service in supporting the partnership between Māori and the Crown.  
(excerpt taken from Factsheet 1)

# Bibliography

- Alkin, M. (2004). Evaluation roots. In Alkin, M. C. *Evaluation roots*, 374-380. Thousand Oaks, CA: Sage.
- Alkin, M.C. & Christie, C.A. (2005). Unravelling theorists' evaluation reality. *New Directions for Evaluation*, 106, 111–28.
- Altmayer, C. (2006). Moving to performance-based management. *Government Finance Review*, 22(3), 8.
- Alvesson, M. & Skoldberg, K. (2009). *Reflexive methodology: New vistas for qualitative research*. Thousand Oaks, CN: Sage.
- Andersen, B., Henriksen, B., & Aarseth, W. (2006). Holistic performance management: an integrated framework. *International Journal of Productivity and Performance Management*, 51(1), 61–78.
- Andersen, N. A. (2008). *Partnerships: Machines of possibility*. Bristol, UK: Policy Press.
- Anil, A. (Ed.). (2016). *Managing big data integration in the public sector*. Hershey, PA: IGI Global.
- Appleton-Dyer, S., Clinton, J., Carswell, P., & McNeill, R. (2012). Understanding evaluation influence within public sector partnerships: A conceptual model. *American Journal of Evaluation*, 33(4), 532–546.
- Archer, M., Bhaskar, R., Collier, A., Lawson, T., & Norrie, A. (Eds). (1998). *Critical realism*. London. UK: Routledge.
- Argyris, C. (1990a). *Organizational learning*. Pittsburgh: PA: Springer.
- Argyris, C. (1990b). *Overcoming organizational defenses: Facilitating organizational learning*. Needham Heights, MA: Allyn & Bacon.
- Argyris, C. (1997). Learning and teaching: A theory of action perspective. *Journal of Management Education*, 21(1), 9-26.
- Argyris, C. (1999). *On organizational learning*. (2<sup>nd</sup> ed.). Maden, MA: Blackwell.
- Argyris, C., & Schön, D. (1974). *Theory in practice*. San Francisco, CA: Jossey-Bass.
- Armon, J. (2007). Aid, politics and development: A donor perspective. *Development Policy Review*, 25(5), 653–656.
- Armstrong, G. (2009). *Making RBM simpler: A practical approach to results-based management*. <https://www.rbmtraining.com/>
- Ashour, A. S. (2005). *Good governance for development in the Arab Countries: Development-driven public governance. Reform in the Arab Countries*. Dead Sea, Jordan: UNDP.
- Aucoin, P. (2005). *Decision-making in government: The role of program evaluation*. Treasury Board of Canada Secretariat. <https://www.tbs-sct.gc.ca/cee/tools-outils/aucoin-eng.asp>
- Aucoin, P. (2012). New political governance in Westminster systems: Impartial public administration and management performance at risk. *Governance*, 25.
- Aucoin, P., & Heintzmann, R. (2000). The dialectics of accountability for performance in public management reform. *International Review of Administrative Sciences*, 66(1), 45-55.



- Aucoin, P., & Jarvis, M. (2004). *Results-based reporting: Smart practices for improving public accountability*. Dalhousie University, Canada: Canada School of Public Service.
- Auditor General of Canada. (2002). Modernizing accountability in the public sector. *Report of the Auditor General of Canada to the House of Commons*, Ch. 9. Ottawa, Canada: Auditor General of Canada.
- Australian Agency for International Development (AusAID). (2000). *AusGUIDE: Stage 4: Mobilisation, implementation and monitoring*. <https://www.aid.gov.au>
- Australian Agency for International Development (AusAID). (2003). AusAID baseline study guidelines. <https://www.aid.gov.au>
- Australian Agency for International Development (AusAID). (2005). *Activity design: The logical framework approach*. <https://www.aid.gov.au>
- Australian Agency for International Development (AusAID). (2006). *FAQ on Paris Declaration and Aid Effectiveness*. <https://www.aid.gov.au>
- Australian Agency for International Development (AusAID). (2007). *Policy: Performance assessment and evaluation*. <https://www.aid.gov.au>
- Australian Government. (2011). *Independent Review of Aid Effectiveness*. Canberra, Australia: Commonwealth of Australia.
- Australian Office of Aid Effectiveness. (2008). *Mutual accountability: 'Orphan' principle of the Paris Declaration*. <https://www.ode.aid.gov.au/resources/index.html>
- Ba Tall, K. (2009). The role of national, regional and international organizations in strengthening country-led M&E systems. In *Country-led M&E systems: Better evidence, better policies, better development results* (119–134). New York, NY: UNICEF.
- Bamberger, M. (2000). The evaluation of international development programs: A view from the front. *American Journal of Evaluation*, 21(1), 95–102.
- Bamberger, M. (2010). Institutionalizing impact evaluation. A key element in strengthening country-led M&E systems. In *From policies to results: Developing capacities for country M&E systems* (pp. 89–216). New York, NY: UNICEF.
- Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: working under budget, time, data, and political constraints*. Thousand Oaks, CA: Sage.
- Barclay-Kerr, H. (2006). Waka – canoes - waka in New Zealand. *Te Ara - the encyclopaedia of New Zealand*. <https://www.TeAra.govt.nz/en/waka-canoes/page-2>
- Barrados, M., & Mayne, J. (2003). Can public sector organisations learn? *OECD Journal on Budgeting*, 3(3), 87–103.
- Barrett, A. M. (2011). An education millennium development goal for quality: complexity and democracy. *Compare: A Journal of Comparative and International Education*, 41(1), 145–148.
- Baser, H., & Morgan, P. (2002). *Harmonising the provision of technical assistance: Finding the right balance and avoiding the new religion*. ECDPM Discussion Paper, 36. Maastricht, Netherlands: ECDPM.
- Bassey, M. (1999). *Case study research in educational settings*. Buckingham, UK: Open University Press.
- Bazeley, P. (2007). *Qualitative data analysis with NVivo*. London, UK: Sage.

- Beloe, T. (2005). *Moving towards local definitions of aid effectiveness: How is the OECD DAC Paris Declaration being interpreted in DFID South East Asia partner countries?* Bangkok, Thailand: Department for International Development.
- Bernstein, D. J. (1999). Comments on Perrin's 'effective use and misuse of performance measurement'. *American Journal of Evaluation*, 20(1), 9.
- Betts, J., & Wedgwood, H. (2011). Effective institutions and good governance for development. *Evaluation Insights*, 4.  
<https://www.oecd.org/derec/unitedkingdom/50313780.pdf>
- Bezzi, C. (2006). Evaluation pragmatics. *Evaluation*, 12(1), 56–76.
- Bhaskar, R. (1979/1998). *The possibility of naturalism: A philosophical critique of the contemporary social sciences*. London, UK: Routledge.
- Bhaskar, R. (1989). *Reclaiming reality*. London, UK: Verso.
- Binnendijk, A. (2000). *Results-based management in the development cooperation agencies: A review of experience. Background report*. Paris, France: OECD-Working Party on Aid Evaluation.  
<https://www.oecd.org/dataoecd/17/1/1886527.pdf>
- Binnendijk, A. (2001). *Results-based management in donor agencies*: DAC Working Party on Aid Evaluation. Paris, France: OECD.
- Bishop, R. (2005). Freeing ourselves from neo-colonial domination in research: A Māori approach to creating knowledge. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Handbook of Quality Research* (pp.109–138). Thousand Oaks, CN: Sage.
- 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.
- Brinkerhoff, D. W. (1996). Process perspectives on policy change: Highlighting implementation. *World Development*, 24(9), 1395–1401.
- Brinkerhoff, R. O. (2002). *The success case method: find out quickly what's working and what's not*. San Francisco, CA: Berrett-Koehler.
- Brinkerhoff, R. O. (2009). *Training the success case method: A strategic evaluation approach to increasing the value and effect of training*. Conference proceedings. <https://www.semanticscholar.org/paper/Training-The-Success-Case-Method%3A-A-Strategic-to-of-Brinkerhoff/8d03c9b5d57f0bc9d3c945983b6891085766b96a>
- British Columbia Office of the Auditor General. (1997). *Reporting performance information paper* presented at the Symposium on Public Sector Accountability Reporting and Auditing Issues. Vancouver, Canada.
- Buffardi, A. L. (2016). *When theory meets reality. Assumptions, feasibility and implications of a complexity-informed approach. A Methods Lab publication*. London, UK: Overseas Development Institute.
- Buffardi, A. L., Mason, P., Hutchings, C., & Sharp, S. (2019). *Monitoring and learning for country-level portfolio decision-making and adaptation. Briefing Note*. London, UK: Overseas Development Institute.
- Burdett, M. (2002). *National development frameworks, local accountability and democracy in Honduras*. Paper presented at the Democratising Development: Deepening Social Accountability through PRSPS. Washington, DC.
- Burgess, R.G. (1984). *In the Field: An introduction to field research*. New York, NY: Allen & Unwin.

- Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organizational analysis*. Aldershot, UK: Gower.
- Busan High Level Forum on Aid Effectiveness. (2011). *Proceedings. Fourth High Level Forum on Aid Effectiveness (HLF-4)*.  
<https://www.oecd.org/dac/effectiveness/HLF4%20proceedings%20entire%20doc%20for%20web.pdf>
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In *Experimental and quasi-experimental designs for research* (pp. 1–34). Boston, MA: Houghton Mifflin.
- Canadian Evaluation Society. (2010). *Competencies for Canadian evaluation practice*.  
<https://evaluationcanada.ca/competencies-canadian-evaluators>
- Canadian International Development Agency. (2007). *Results-based management*.  
<https://www.acdi-cida.gc.ca/CIDA>
- Capacity Development for Development Effectiveness Facility. (2010). *Implementing the Paris Declaration and the Accra Agenda for Action in Asia and the Pacific*. Manila, Philippines: Asian Development Bank <https://www.acdi-cida.gc.ca/CIDA>
- Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*. 15(2), 195-207.  
<https://doi:10.3102/01623737015002195>
- Carlson, T., Moewaka Barnes, H., & McCreanor, T. (2017). Kaupapa Māori evaluation: A collaborative journey. *Evaluation Matters—He Take Tō Te Aromatawai: Online New Zealand Council for Educational Research*.  
<https://www.nzcer.org.nz/nzcerpress/evaluation-matters/articles/kaupapa-m-ori-evaluation-collaborative-journey>
- Carlsson, S. A. (2011). *Critical Realist Information Systems Research in Action*. Conference Proceedings. Berlin, Germany. <https://opendl.ifip-tc6.org/db/conf/ifip8-2/ifip8-2-2011/Carlsson11.pdf>
- Castles, F.G. (2001). On the political economy of recent public sector development. *Journal of European Social Policy*, 11(3), 195–211.
- Castro, M. F. (2008). *Insider insights: Building a results-based management and evaluation system in Colombia*. ECD Working Paper Series; 18. Washington, DC: The World Bank.
- Central Agencies Steering Group. (2002). *Managing for outcomes: Output plans guidance for departments*. Letter to the Chief Executive. Steering Group Managing for Outcomes Roll-out 2003/04. Wellington, New Zealand: New Zealand Government.
- Chakravartty, P. (2007). Governance without politics: Civil society, development and the postcolonial state. *International Journal of Communication*, 2007(1), 297–317.
- Chang, H-J. (2003). (Ed.). *Rethinking development economics*. London, UK: Anthem.
- Chambers, R. (1983). *Rural development: putting the last first*. Harlow, UK: Prentice Hall.
- Chambers, R. (2010). *Paradigms, poverty and adaptive pluralism*. IDS Working Papers, 2010 (344), 1-57. [https://doi:10.1111/j.2040-0209.2010.00344\\_2.x](https://doi:10.1111/j.2040-0209.2010.00344_2.x)
- Chand, S. (2010). Overview. T. Thomas, & L. Duncan, (Eds.). In *Papua New Guinea's development performance, 2005-2008* (pp. 1-16). Port Moresby, Papua New Guinea: National Research Institute.

- Chelimsky, E., & Shadish, W. R. (Eds.). (1997). *Evaluation for the 21st century: a handbook*. Thousand Oaks, CA: Sage.
- Chen, H. (1990). Issues in constructing program theory. *New Directions for Program Evaluation*, 1990(47), 7–18.
- Chen, H. (1990). *Theory-driven evaluations*. Thousand Oaks, CA: Sage Newbury Park
- Chen, H., & Rossi, P. H. (1980). The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social Forces*, 59 (1), 106–122.
- Chimhowu, A. O., Hulme, D., & Munro, L. T. (2019). The ‘new’ national development planning and global development goals: Processes and partnerships. *World Development*, 120, 76–89.
- Chouinard, J. A., Cousins, J. B. (2009). A review and synthesis of current research on cross- cultural evaluation. *American Journal of Evaluation*, 30, 457–494.
- Chouinard, J. A. (2013). The practice of evaluation in public sector contexts: A response. *American Journal of Evaluation*, 34, 266–269.
- Christiansen, K. & Rogerson, A. (2005) *Is the current aid architecture ‘fit for purpose’?* London, UK: Overseas Development Institute.
- Christensen, T. (Ed.), Lægreid, P. (2007). *Transcending new public management*. London, UK: Routledge.
- Christie, C. A., & Alkin, M. C. (2003). The user-oriented evaluator’s role in formulating a program theory: Using a theory-driven approach. *American Journal of Evaluation*, 24(3), 373-385.
- Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, 34(3), 31-135.
- Clegg, S., & McNulty, K. (2002). Partnership working in delivering social inclusion: organizational and gender dynamics. *Journal of Education Policy*, 17(5), 587–601.
- Conlin, S. & Stirrat, R. (2008). Current challenges in development evaluation. *Evaluation*, 14(193).
- Cook, A. L. (2004). *Managing for outcomes in the New Zealand public management system*. Wellington, New Zealand: The Treasury.
- Corre, G., Mackie, J., & Trenner. (2008). *Current dilemmas in aid architecture. Actors & instruments, aid orphans and climate change. Policy Management Report 16*. European Centre for Development Policy Management.  
<https://ecdpm.org/publications/dilemmas-aid-actors-instruments-aid-orphans-climate-change/>
- Cram, F. (2006). Talking Ourselves UP. *AlterNative: An International Journal of Indigenous Peoples*, 2(1), 28-43.
- Cram, F. (2009). Maintaining indigenous voices. In: Mertens, D., Ginsberg, P. (Ed). *The handbook of social research ethics*, 308–322. Los Angeles, CN: Sage.
- Cram, F. (2016). Lessons on decolonizing evaluation from Kaupapa Māori evaluation. *Canadian Journal of Program Evaluation*, 30(3), 296–312.
- Cram, A., Hopson, R., Powell, M., Williams, A., & Kaul, A. (2019). Challenges and possibilities in developing a programme-theory model through stakeholder engagement, dialogue, and reflection. *Evaluation Matters – He Take Tō Te Aromatawai: Online First*.  
[https://www.nzcer.org.nz/system/files/journals/evaluation-matters/downloads/NZCER\\_Evaluation\\_Matters\\_CramEtAl.pdf](https://www.nzcer.org.nz/system/files/journals/evaluation-matters/downloads/NZCER_Evaluation_Matters_CramEtAl.pdf)

- Cram, F., Pipi, K., & Paipa, K. (2018). Kaupapa Māori evaluation in Aotearoa New Zealand. *New Directions for Evaluation*, 2018(159), 63–77.
- Crawford, G. (2003). Promoting democracy from without – Learning from within (part 1). *Democratization*, 10(1), 77–98.
- Creswell, J. W. (2003a). *Research design: qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2003b). Advanced mixed methods research designs. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioural Research*. London, UK: Sage.
- Crocombe, R. G. (2001). *The South Pacific* (rewritten, updated & expanded ed.). Suva, Fiji: Institute of Pacific Studies, University of the South Pacific.
- Crotty, M. (ed.) (1998). *The foundations of social research: meaning and perspective in the research process*. Thousand Oaks, CA: Sage.
- Cunningham, C. (2000). A framework for addressing Māori knowledge in research, science and technology. *Pacific Health Dialog*, 7(1), 62-69.
- Cunningham, C. (2003). Application of the TPK guidelines for conducting evaluation research with Māori. In N. Lunt, C Davidson, & K. McKegg, K. (Eds), *Evaluating policy and practice: A New Zealand reader*. Auckland, New Zealand: Pearson Education.
- Dabelstein, N., & Patton, M.Q. (2013). The Paris declaration on aid effectiveness: History and significance. *The Canadian Journal of Program Evaluation*, Vol. 27 No. 3, 19–36. Canadian Evaluation Society.
- Dahler-Larsen, P. (2012). *The evaluation society*. Stanford, CN: Stanford University Press.
- Dahler-Larsen, P., & Boodhoo, A. (2019). Evaluation culture and good governance: Is there a link? *Evaluation*, 25(3), 277-293.
- Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhusser, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*. 10, 49.
- Dalziel, P. & Saunders, C. (2014). *Wellbeing economics: A policy framework for New Zealand*. Christchurch, New Zealand: Lincoln University.
- Davies, J. S. (2007). The limits of partnership: An exit-action strategy for local democratic inclusion. *Political Studies*, 55(4), 779-800.
- Davies, R. (2004). Scale, complexity and the representation of theories of change. *Evaluation*, 10(1), 101–121.
- Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage.
- De Gennaro, I. (2012). *Value: Sources and readings on a key concept of the globalized world*. Leiden, Netherlands: Brill.
- Denzin, N. K., & Giardina, M. D. (2010). *Qualitative inquiry and human rights*. Walnut Creek, CA: Left Coast Press.
- Department of International Development (DFID). (2009). *Building the evidence to reduce poverty: The UK's policy on evaluation for international development*. London, UK: UK Government.
- Department for International Development. (2006). *Eliminating world poverty: Making governance work for the poor*. London, UK: UK Government.

- Department for International Development. (2005). *Guidance on evaluation and review for DFID staff*. London, UK: UK Government.  
<https://www.dfid.gov.uk/aboutdfid/performance/files/guidance-evaluation.pdf>
- Department for International Development. (2009). *Eliminating world poverty: Building our common future*. London, UK: UK Government.
- Department of Families. (2003). *Queensland families: Future directions strategy – a smart state initiative. Evaluation framework*. Queensland, Australia: Queensland Government.
- Department of National Planning and Monitoring (DNPM). (2014). *National strategy for responsible sustainable development for Papua New Guinea – A paradigm shift – addendum to the Development Strategic Plan, 2010–2030*. Port Moresby, Papua New Guinea: Department of National Planning and Monitoring.
- Department of National Planning and Monitoring (DNPM). (2016). *Monitoring and evaluation framework*. Port Moresby, Papua New Guinea: Department of National Planning and Monitoring.
- Department of the Auditor General of Canada. (n.d.). *Auditor General's report in systems under development*. <https://www.oag-bvg.gc.ca/domino/other.nsf/html/report.html>
- Department of Treasury. (2009). *Vision 2050*. Port Moresby, Papua New Guinea: PNG Government.
- Dey, I. (1999). *Grounding grounded theory: Guidelines for qualitative inquiry*. London, UK: Academic Press.
- Dobell, R. (2003). *The role of government and the government's role in evaluating government: Insider information and outsider beliefs*. Panel on the Role of Government Research Paper. Vancouver, Canada: University of Victoria.
- Donaldson, S. I. (2003). Diverse visions for evaluation in the new millennium: Should we integrate or embrace diversity? In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 3–16). Mahwah, NJ: Erlbaum.
- Donaldson, S. I., Christie, C. A., & Mark, M. M. (Eds.). (2009). *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.
- Donaldson, S. I., & Scriven, M. (Eds.). (2003). *Evaluating social programs and problems: Visions for the new millennium*. Mahwah, NJ: Lawrence Erlbaum.
- Dormer, R., & Gill, D. (2010). Managing for performance in New Zealand's public service: A loosely coupled framework. *Measuring Business Excellence*, 14(1), 43-59.
- Dormer, R., & Ward, S. (2018). *Accountability and public governance in New Zealand. Working Paper No. 117*. Wellington, New Zealand: Victoria University of Wellington.
- Droop, J., Insenman, P., & Mlalazi, B. (2008). *Paris Declaration on aid effectiveness: Study of existing mechanisms to promote mutual accountability (between donors and partner countries at the international level. Background paper: Review of individual mechanisms*. Paris, France: OECD.
- Drucker, P. (1955). *The practice of management*. Oxford, UK: Elsevier.
- Dryzek J. S. (2000) *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford, UK: Oxford University Press.

- Dugarova, E. (2015). *Social inclusion, poverty eradication and the 2030 agenda for sustainable development*, UNRISD Working Paper, No. 2015-15. Geneva, Switzerland: United Nations Research Institute for Social Development (UNRISD). <https://hdl.handle.net/10419/148736>
- Duignan, P. (2004). *The use of formative evaluation by government agencies*. <https://www.strategicvaluation.info/se/documents/121pdf.html>.
- Duignan, P. (2008). *What added value can evaluators bring to governance, development and progress through policy-making? The role of large visualized outcomes models in policy-making*. Conference paper. 8th European Evaluation Society Conference, Lisbon, Portugal, October 2008.
- Duignan, P. (2009). Introduction to outcomes theory. *Outcomes theory knowledge base article no. 218*. <https://knol.google.com/k/paul-duignan-phd/introduction-to-outcomes-theory/2m7zd68aaz774/3>
- Duignan, P. (2010). Outcomes-focused visual strategic planning for public and third sector organizations. *Outcomes theory knowledge base, article no. 287*. <https://knol.google.com/k/paul-duignan-phd/duignan-s-outcomes-focused>
- Duignan, P. Building social policy evaluation capacity. *Social Policy Journal of New Zealand*, 19.
- Durant, R. F. (1999). The political economy of results-oriented management in the "neo-administrative state": Lessons from the MCDHHS experience. *The American Review of Public Administration* 1999, 29, 307–331.
- Durie, M. (2003). *Keynote address: Is there a distinctive Māori psychology?* Proceedings of the National Māori Graduates of Psychology Symposium 2002: Making a difference. Proceedings of a symposium hosted by the Māori & **Psychology Research Unit at the University of Waikato, Hamilton, 29-30 November 2002 (pp.19-25)**. Hamilton, New Zealand: Māori and Psychology Research Unit, University of Waikato. <https://hdl.handle.net/10289/849>.
- Durie, M. (2004). *Public sector reform, indigeneity, and the goals of Māori development*. Paper presented at the Commonwealth Advanced Seminar, Wellington, New Zealand.
- Durie, M. (2006). *Measuring Māori wellbeing*. New Zealand Treasury guest lecture series. Wellington, New Zealand: The Treasury.
- Durie, M., Fitzgerald, M., Kingi, T., McKinley, S., & Stevenson, B. (2002). *Māori specific outcomes and indicators: A report prepared for Te Puni Kōkiri*. Wellington, New Zealand: Te Puni Kōkiri
- Durie, M., Fitzgerald, E., Kingi, T. K., McKinley, S., & Stevenson, B. (2003). *Monitoring Māori progress Te Ngahuru: A Māori outcome framework*. Paper presented at the Social Policy Research and Evaluation Conference, Wellington, New Zealand.
- Dwivedi O. P., Khator R., & Nef J. (2007). The ethics of development: Management in a global era. In *Managing Development in a Global Context*. London, UK: Palgrave Macmillan.
- Dugarova, E. (2015). *Social inclusion, poverty eradication and the 2030 agenda for sustainable development*. Geneva, Switzerland: United Nations Research Institute for Social Development (UNRISD). <https://hdl.handle.net/10419/148736>
- Earle, L. (2002). *Lost in the matrix: The logframe and the local picture*. Paper presented at the INTRAC's 5th Evaluation Conference: Measurement, Management and Accountability, The Hague, Netherlands.

- Easton, G. (2010). Critical realism in case study research. *Industrial Marketing Management*, 39(1), 118-128.
- Entwistle, J. K. (2007). *Results-based national development strategies: Assessment and challenges ahead*. Conference Proceedings. The World Bank.
- Epstein, J., & Olsen, R. T. (1996). Lessons learned by state and local governments. *The Public Manager*, Fall, 41-44.
- Eyben, R. (2008). *Power, mutual accountability and responsibility in the practice of international aid: A relational approach*. Working Paper No. 305. Brighton, UK: Institute of Development Studies.
- Farrelly, T. & Nabobo-Babat U. (2014). Talanoa as empathic apprenticeship. *Asia Pacific Viewpoint*, Vol. 55(3), 319-330.
- Fear, H., & Barnett, P. (2003). Holding fast: The experience of collaboration in a competitive environment. *Health Promotions International*, 18(1), 6-14.
- Feinstein, O. N. (2017). *Evaluation and poverty reduction* (1st ed.). London, UK: Taylor and Francis.
- Ferguson, J. (1990). *The anti-politics machine: "development," depoliticization, and bureaucratic power in Lesotho*. Cambridge, UK: Cambridge University Press.
- Ferguson, J. (2002). Spatializing states: Towards an ethnography of neoliberal governmentality. *American Ethnologist*, 29(4), 981-1002.
- Ferguson, N. (2003). British imperialism revised: The costs and benefits of 'Anglobalization'. *Historically Speaking*, 4, 21-27.  
<https://doi:10.1353/hsp.2003.0063>
- Ferreira, A., Azevedo, G., Oliveira, J., & Marques, R. (2016). *Global perspectives on risk management and accounting in the public sector*. Information Science Reference, an imprint of IGI Global.
- Fetterman, D. F. (2004). Branching out or standing on a limb: Looking to our roots for Insight. In M. C. Alkin (Ed.), *Evaluation roots. Tracing the theorists' views and influences*. Thousand Oaks, CA: Sage.
- Fetterman, D. M. & Wandersman, A. (Eds.). (2005). *Empowerment evaluation principles in practice*. New York, NY: Guilford Press.
- Fforde, A. (2013). *Understanding development economics: Its challenge to development studies*. London, UK: Routledge.
- Fforde, A. (2017). Development in the early years and the facts of underdevelopment since WWII. In *Reinventing Development*. [https://link-springer-com.ezproxy.massey.ac.nz/chapter/10.1007/978-3-319-50227-4\\_3](https://link-springer-com.ezproxy.massey.ac.nz/chapter/10.1007/978-3-319-50227-4_3)
- Forss, K., & Bandstein, S. (2008). Evidence-based evaluation of development cooperation: Possible? Feasible? Desirable? *IDS Bulletin* 39(1). London, UK: Institute of Development Studies.
- Forss, K., Kruse, S.-E., Taut, S., & Tenden, E. (2006). Chasing a ghost? An essay on participatory evaluation and capacity development. *Evaluation*, 12(1), 128-144.
- Frot, E., & Santiso, J. (2010). *Crushed aid: Fragmentation in sectoral aid*. SITE Working Paper No. 6, 2009. Stockholm, Sweden: Stockholm Institute of Transition Economics.
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: effective use of theories of change and logic models* (1st ed.). San Francisco, CA: Jossey-Bass.
- Gaspar, D. (1999). *Evaluating the "logical framework approach": towards learning-oriented development evaluation*. The Hague, Netherlands: Institute of Social Studies, Netherlands.



- Gasper, D. (2001). *Logical frameworks: Problems and pitfalls*.  
<https://winelands.sun.ac.za/2001/Papers/Gasper,%20Des.htm>
- Geddes, M., Mensah, E. J., Natali, L., & Quartey, P. (2009). *OECD DAC – PDE thematic study on untied aid: Ghana country study*. Paris, France: OECD.
- Geoghegan, T., D’Errico, S., Acuña, M. G., El-Saddik, K., Lucks D., Ocampo A. & Piergallini I. (2019). *Evaluating sustainable development: how the 2030 Agenda can help*. The International Institute for Environment and Development (IIED).  
<https://pubs.iied.org/17713IIED/>
- George, A. L. and Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.
- Giddens, A. (1998). *The third way: The renewal of social democracy*. Cambridge, UK: Polity Press.
- Gill, D. (2008). Managing for performance in New Zealand: The search for the “holy grail”? In *Holy Grail or Achievable Quest? International Perspectives on Public Sector Performance Management*. Canada: KPMG.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Mill Valley, CA: Sociology Press.
- Goetz, A. M., & Jenkins, R. (2005). *Reinventing accountability: making democracy work for human development*. New York, NY: Palgrave Macmillan.
- Goldfinch, S. (2014). Evaluating public sector reform in New Zealand: Have the benefits been oversold? *Asian Journal of Public Administration*, 20, 203–232.
- Government of Papua New Guinea. (2000). *Memorandum on economic and financial policies of the Government of Papua New Guinea*.
- Grant, R. M. (1998). *Contemporary strategy analysis* (3rd ed.). Malden, MA: Blackwell
- Gray, A. (2002). *Integrated service delivery and regional co-ordination: a literature review prepared as part of the review of the centre – regional co-ordination workstream*. Wellington, New Zealand: Gray Matter Research.
- Gray, & Bray. (2019). *Evaluation in the Australian public service: current state of play, some issues and future directions*. An ANZSOG research paper for the Australian Public Service Review Panel. The Australia and New Zealand School of Government (ANZSOG).
- Grey, A., Jenkins, B., Leeuw, F., & Mayne, J. (2003). *Collaboration in public services: The challenge for evaluation*. With foreword by Desautels, the Auditor General of Canada (1991–2001). New Brunswick, Canada: Transaction.
- Grindle, M. S. (2007). Good enough governance revisited. *Development Policy Review*, 25(5), 533-574.
- Grindle, M. S., & Hilderbrand, M. E. (1995). Building sustainable capacity in the public sector: what can be done? *Public Administration & Development*, 15(5), 441–463.
- Grove, J., Kibel, B., & Haas, T. (2007). EvaluLEAD: An open-systems perspective on evaluating leadership development. In K. Hannum, J. Martineau, & C. Reinelt (Eds.), *Handbook of leadership development evaluation*. San Francisco, CA: Jossey-Bass.
- Guba, Y., & Lincoln, E. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage
- Guerra-Lopez, I. J. (2008). *Performance Evaluation* (1<sup>st</sup> ed.). San Francisco, CA: Jossey-Bass.
- Gupta, A. I., & Sharma, A. (2006). Globalization and postcolonial states. *Current Anthropology*, 47(2).

- Habraken, R., Schulpen, L., & Hoebink, P. (2017). Putting promises into practice: The New Aid Architecture in Uganda. *Development Policy Review*, 35(6), 779-795.
- Haggard, S., & Kaufman, R. R. (1994). The challenges of consolidation. *Journal of Democracy* 5(4), 5-16.
- Hale, J. A. (2004). *Performance-based management: what every manager should do to get results*. San Francisco, CA: Pfeiffer Wiley.
- Hammersley, M. (2009). Why critical realism fails to justify critical social research. *Methodological Innovations Online*, 4(2), 1-11.
- Hansen, M. B., & Vedung, E. (2010). Theory-based stakeholder evaluation. *American Journal of Evaluation*, 31(3), 295–313.
- Harré, R. (1978). *The augmentation of psychology*, *Zygon®*, 13(1), 34-52.
- Hatry, H., Cowan, J., & Hendricks, M. (2004). *Analysing outcome information: Getting the most from data*. Washington, DC: The Urban Institute.
- Hatry, H., & Lampkin, L. (2003). *Key steps on outcome management*. Washington, DC: The Urban Institute.
- Hawkins, A. J. (2020). Program logic foundations: Putting the logic back into program logic. *Journal of MultiDisciplinary Evaluation*, 16(37), 38-57.
- Hayward, J. (2012). Principles of the Treaty of Waitangi – ngā mātāpono o te tiriti'. *Te Ara - the Encyclopaedia of New Zealand*.  
<https://www.TeAra.govt.nz/en/principles-of-the-treaty-of-waitangi-nga-mataponono-o-te-tiriti/print>
- Health Quality & Safety Commission. (2019). *He matapihi ki te kounga o ngā manaakitanga ā-hauora o Aotearoa: A window on the quality of Aotearoa New Zealand's Health Care*. [https://www.hqsc.govt.nz/assets/Health-Quality-Evaluation/PR/Window\\_2019\\_web\\_final.pdf](https://www.hqsc.govt.nz/assets/Health-Quality-Evaluation/PR/Window_2019_web_final.pdf)
- H.M. Treasury. (2013). *The green book: Appraisal and evaluation in central government*.
- Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London, UK: McGraw-Hill.
- Holmes, A. G. D. (2020). Researcher positionality – a consideration of its influence and place in qualitative research – a new researcher guide. *Shanlax International Journal of Education*, 8(4), 1-10.
- Horner, R. (2019). Towards a new paradigm of global development? Beyond the limits of international development. *Progress in Human Geography*, 0(0).
- Horner, R., & Hulme, D. (2019). Global development, converging divergence and development studies: A rejoinder. *Development and Change*, 50(2), 495-510.
- House, E. R. (2001). Responsive evaluation (and its influence on deliberative democratic evaluation). *New Directions for Evaluation*, 2001(92), 23.
- Howlett, M. (2002). Understanding national administrative styles and their impact upon administrative reform: A neo-institutional model and analysis. *Policy and Society*, 21, 1-24.
- Huemer, W., & Centi, B. (2009). *Values and ontology: problems and perspectives*. Berlin, Germany: De Gruyter.
- Hughes, O. E. (2003). *Public management and administration: an introduction (3rd ed.)*. London, UK: Springer Nature.

- Hughes, O. E. (2009). International handbook of practice-based performance management. In E. A. Julnes (Ed.), *Performance: A New Public Management Perspective* (p. 1). Thousand Oaks, CA: Sage.
- Hughes, P., & Smart, J. (2012). You say you want a revolution: the next stage of public sector reform in New Zealand. *Policy Quarterly*, 8(1).
- Hulme, D., Savoia, A., & Sen, K. (2015). Governance as a global development goal? Setting, measuring and monitoring the post-2015 development agenda. *Global Policy*, 6 (2), 85-96.
- Humpage, L. (2005). Experimenting with a 'whole of government' approach. *Policy Studies*, 26(1), 47-66.
- Hummelbrunner, R., & Jones, H. (2013). *A guide to managing in the face of complexity*. ODI Background Note. London: Overseas Development Institute.
- Hunt, T. (n.d.). *Innovations in accountability: Federal government directions in strengthening accountability and the central role of evaluation*. New Brunswick, Canada: Treasury Board of Canada.
- Hurworth, R. (2008). Program clarification: an overview and resources for evaluability assessment, program theory and program logic. *Evaluation Journal of Australasia*, 8(2), 42–48.
- Innes, J., & Booher, D. (2018). *Planning with complexity: An introduction to collaborative rationality for public policy* (2<sup>nd</sup> ed.). Abingdon, UK: Routledge.
- Jackson, E. T. (1998). Indicators of change: *Results-based management and participatory evaluation*. *Knowledge Shared: Participatory Evaluation in Development Cooperation*. West Hartford, CT: Kumarian Press.
- Jackson, P. (1993). Public service performance evaluation: A strategic perspective. *Public Money & Management*, 13, 9–14.
- James, C. (1992). *New Territory: The Transformation of New Zealand, 1984-1992*. Wellington, New Zealand: Bridget Williams Books.
- Jones, H. M.F. (2000). Partnerships: a common-sense approach to inclusion? In: *Researching "inclusion." Papers from the Annual Conference of the Standing Conference on University Teaching and Research in the Education of Adults 2000*. SCUTREA, (pp. 182-188). Boston, UK.  
<https://eprints.hud.ac.uk/id/eprint/10608>
- Joyce, P. G. (1997). Using performance measures for budgeting: A new beat, or is it the same old tune? *New Directions for Evaluation*, Fall (75), 45–61.
- Kalimullah, N., Ashraf, K., & Ashaduzzaman, N. (2019). *New public management: Emergence and principles*, (1) 1–22.
- King, J. M. (2007). *Eke ki runga i te waka: The use of dominant metaphors by newly fluent Māori speakers in historical perspective*. PhD thesis. Christchurch, New Zealand: University of Canterbury.
- Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, July–August 2007, 150–170.
- Kaufmann, D., Kraay, A. (2002). *Growth without governance*. *Policy, research working paper series; no. WPS 2928*. Washington, DC: World Bank.  
<http://documents.worldbank.org/curated/en/811781468766468180/Growth-without-governance>
- Kaufmann, D., & Kraay, A. (2005). *Governance indicators: Where are we, where should we be going? Policy Research Working Paper, 4370*. Washington, DC: The World Bank.

- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). Response to 'what do the worldwide governance indicators measure?' *Journal of the European Association of Development Research and Training Institutes [EADI]*, 22(1), 55–58.
- Kawakami, A., Aton, K., Cram, F., Lai, M., & Porima, L. (2007). Improving the practice of evaluation through indigenous values and methods: Decolonising evaluation practice – Returning the gaze from Hawai'i and Aotearoa. *Hulili: Multidisciplinary Research on Hawaiian Well-Being*, 4 (1).
- Kerr, S. (2012). Kaupapa Māori theory-based evaluation. *Evaluation Journal of Australasia*, 12(1) 6–18.
- Khazanchi, D., & Owens, D. (2018). *From strategic intent to implementation: How information technology initiatives take shape in organizations*. Paper presented at the *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii.
- Killen, B., & Rogerson, A. (2010). *Global governance for international development: Who's in charge?* OECD Development Brief. Paris, France: OECD.
- Kindornay, S., & Morton, B. (2009). *Development effectiveness: towards new understandings. Issues Brief: Development Cooperation Series*, 1–6. The North South Institute.
- Kingi, T., & Durie, M.H. (2000). *Hua Oranga: A Māori measure of mental health outcomes*. Palmerston North, New Zealand: Te Pūmanawa Hauora, School of Māori Studies, Massey University.
- Kirkhart, K. E. (2005). Through a cultural lens: Reflections on validity and theory in evaluation (Chapter 3), In S. Hood, R. Hopson, H. Frierson (Eds.), *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding in evaluative theory and practice* (pp. 21-39). Greenwich, CT: Information Age.
- Kotvojs, F. (2006). *Contribution analysis – A new approach to evaluation in international development*. Paper presented at the Australasian Evaluation Conference 2006, Darwin, Australia.
- Knack, S. (2013). Aid and donor trust in recipient country systems. *Journal of Development Economics*, 101, 316-329. doi.org/10.1016/j.jdeveco.2012.12.002
- Kraemer-Mbula, E., & Wamae, W. (2010). *Innovation and the development agenda*. Ottawa, Canada: OECD
- Krznaric, R. (2007). *How change happens: Interdisciplinary perspectives for human development*. Oxfam Research Report, Dublin, UK: Oxfam.
- Kupu Māori (2019). <https://kupu.Māori.nz/kupu/He-waka-eke-noa/>
- Kundin, D. M. (2010). A conceptual framework for how evaluators make everyday practice decisions. *American Journal of Evaluation* 2010 31, 347–362.
- Kusek, J., & Rist, R. (2004). *Ten steps to a results-based M&E system*. Washington, DC: The World Bank.
- Kusek, J. Rist, R. & White, E. (2003). *How will we know the millennium development goal results when we see them? Building a results-based monitoring and evaluation system to give us the answers*. <https://www.managingfordevelopmentresults.org/documents/KusekRistWhitepaper.pdf>
- Lambur, M. T. (1987). Evaluation decision makers. *Journal of Extension*, 25(1).<https://www.joe.org/joe/1987spring/tt2.html>

- Landais-Stamp, P., & Rogers, P. (1989). *Rocking the boat: New Zealand, the United States, and the nuclear-free zone controversy in the 1980s*. Oxford England, New York: Bergan.
- Lanoska, A. (2018). *Socio-economic theories, legacies, and strategies*. New York, NY: Routledge.
- Larson, A. (2018). Evaluation amidst complexity: Eight evaluation questions to explain how complex adaptive systems affect program impact. *Evaluation*, 24(3), 353–362.
- Lavergne, R. (2002). *Results-based management and accountability for enhanced aid effectiveness*. Canadian International Development Agency.
- Lehn, B. (2008). Evaluator's role in accountability relationships: Measurement technician, capacity builder or risk manager? *Evaluation*, 14(3), 323–343.
- Lempert, D. H. (2010). Why government and non-governmental policies and projects fail despite 'evaluations': An indicator to measure whether evaluation systems incorporate the rules of good governance. *Journal of Multidisciplinary Evaluation*, 6(13), 58–108.
- Leonard, K., & Bayley, S. (2008). Improving project, program and policy performance in developing countries through managing for development results. *Evaluation Journal of Australasia*, 8(2).
- Lewis, K. (2007). The paradox of managing for outcomes. *Policy Quarterly*, 3(3), 25–32.
- Liebenthal, A., Feinstein, O. N., & Ingram, G. K. (Eds.). (2004). *Evaluation & development: the partnership dimension*. New Brunswick, NJ: Transaction.
- Lin, J. Y. (2009). *New structural economics: A framework for rethinking development*. The World Bank Research Observer. The World Bank.
- Linder, L., & Brooks, J. (2004). Transforming the public sector. *Outlook*, 2004(3), 76–83.
- Long, N. (1994). *Development sociology*. London, UK: Routledge.
- Lunt, N., Davidson, C., & McKegg, K. (2003). *Evaluating Policy and Practice: A New Zealand Reader*. Auckland New Zealand: Pearsons Education.
- Lusthaus, C., Adrien, M., & Perstinger, M. (1999). Capacity development: Definitions, issues and implications for planning, monitoring and evaluation. *Universalis*, 35.
- Lynch, B. K. (1996). *Language program evaluation: theory and practice*. Cambridge, NY: Cambridge University Press.
- Lynham, S. (2002). The General Method of Theory-Building Research in Applied Disciplines. *Advances in Developing Human Resources*, 4, 221-241. <https://doi:10.1177/1523422302043002>
- Macfarlane, A. (2016). Taiarahia Black (Ed.). Enhancing Mātauranga Māori and global indigenous knowledge. *New Zealand Journal of Educational Studies*, 51(1), 129-32.
- Mackay, K. (1999). *Evaluation capacity development: A diagnostic guide and action framework*. Washington, DC: The World Bank.
- Mackay, K. (2006). *Two generations of performance evaluation and management system in Australia (English)*. Washington, DC: World Bank. <https://documents.worldbank.org/curated/en/595391468768705937/Two-generations-of-performance-evaluation-and-management-system-in-Australia>

- Mackay, K. (2007). *How to build M&E systems to improve government performance: Lessons from Australia, Chile and Colombia*. Presentation to the South African GW M&E Learning Network). Pretoria, South Africa: The World Bank.
- Manning, R., Reveyard, M., & DAC Task Force on Donor Practices. (2003). *Harmonising donor practices for effective aid delivery: good practice papers: a DAC reference document*. Paris, France: OECD.
- Marchal, B., van Belle, S., van Olmen, J., Hoérée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), 192-212.
- Marriott, L., & Sim, D. (2014). *Indicators of inequality for Māori and Pacific People*. Wellington, New Zealand: Victoria Business School.
- Martin-Fagg, R. (1996). *Making sense of the economy*. London, UK: International Thomson Business Press.
- Matheson, A., Shall, A., & Ogeda, M. (2008). *Independent evaluation of Uganda's Poverty Eradication Action Plan (PAEP)*. Final Results and Performance Theme Paper.
- Mathison, S. (2005). *Encyclopaedia of evaluation*. Thousand Oaks, CA: Sage.
- Maxwell, S., & Engel, P. (2003). *European development cooperation to 2010. ECDPM, EADI*. London, UK: Overseas development Institute.
- Maxwell, S., Engel, P., Leiteritz, R. J., Mackie, J., Sunderland, D., & Bettina, W. (2003). *European development co-operation to 2010. Working Paper 219*. London, UK: Overseas Development Institute.
- Maxwell, S., & Mittapali. (2007). *The value of critical realism for qualitative research*. Paper presented at the International Association for Critical Realism, Philadelphia, PA.
- Mayne, J. (1997). Accountability for program performance: A key to effective performance monitoring and reporting. In, J. Mayne, and E. Zapico-Goni, (Eds.), *Monitoring Performance in the Public Sector: Future Directions from International Experience*. New Brunswick, NJ: Transaction Publishers.
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation* 16(1), 1–24.
- Mayne, J. (2004). Reporting on outcomes: Setting performance expectations and telling performance stories. *The Canadian Journal of Program Evaluation*, 19(1), 31–60.
- Mayne, J. (2007a). Best practices in results-based management: A Review of Experience. *A Report for the United Nations Secretariat*, 2 (Annexes).
- Mayne, J. (2007b). Challenges and lessons in implementing results-based management. *Evaluation*, 13(1), 87–109.
- McCarty, A., & Julian, A. (2009). *The developmental effectiveness of untied aid: Evaluation of the implementation of the Paris Declaration and of the 2001 DAC recommendation on untying ODA to the LDCs*. Lao PDR country study.
- McDavid, J., C., & Huse, I. (2007). *Does assuring the credibility of performance reports increase their use?* Vancouver, Canada: School of Public Administration, University of Victoria,
- McDavid, J. C., & Hawthorn, L. R. L. (2006). *Program evaluation and performance measurement: an introduction to practice*. Thousand Oaks, CA: Sage.

- McGuinness, W. (2010). *Report 7: Exploring the shared goals of Māori: Working towards a national sustainable development strategy*. McGuinness Institute. <https://apo.org.au/node/68525>
- McEwan, R. J. (2016). *Opening the black box: how strategy practitioners develop their practices*. PhD thesis. Palmerston North, New Zealand: Massey University.
- McKegg, K. (2003). From margins to mainstream: The importance of people and context in evaluation utilisation. In, N. Lunt, C. Davidson & K. McKegg (Eds.), *Evaluating Policy and Practice: A New Zealand Reader*. Auckland, New Zealand: Pearson Prentice Hall.
- McPhee, I. (2006). *Evaluation and performance audit: Close cousins or distant relatives?* Canberra: Australian National Audit Office.
- Matike Mai Aotearoa New Zealand's constitution: A report on a conversation He Kōtuinga Kōrero mō Te Kaupapa Ture o Aotearoa. (2016). Constitutional Advisory Panel Te Ranga Kaupapa Ture. Wellington, New Zealand: New Zealand Government.
- Medearis J, 2005, Social movements and deliberative democratic theory. *British Journal of Political Science*, 35(1), 53-75.
- Meier, G.M. (2005). *Biography of a subject: An evolution of development economics*. New York, NY: Oxford University Press.
- Menon, S. (2010). A perspective from the United Nations on national ownership and capacity evaluation. In *From policies to results: Developing capacities for country M&E systems* (pp. 117–124). New York, NY: UNICEF.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Ministry for Culture and Heritage. (2017). *Treaty FAQs*. Wellington, New Zealand: Ministry for Culture and Heritage. <https://nzhistory.govt.nz/politics/treaty/treaty-faqs>
- Ministry of Culture and Heritage, (2019). 'Chief Justice declares treaty 'worthless' and a 'simple nullity', *NZ History online*. <https://nzhistory.govt.nz/the-chief-justice-declares-that-the-treaty-of-waitangi-is-worthless-and-a-simple-nullity>
- Ministry of Education. (2005). *Evaluation strategy: Report of the evaluation strategy development project*. Wellington, New Zealand: Ministry of Education.
- Ministry for Pacific Peoples. (2017). *Kapasa: The Pacific policy analysis tool*. Wellington, New Zealand: Ministry for Pacific Peoples.
- Ministry of Health. (2002). *How to define your agency's outcomes*. Wellington, New Zealand: Ministry of Health.
- Minkov, M., & Hofstede, G. (2012). Is national culture a meaningful concept? Cultural values delineate homogeneous national clusters of in-country regions. *Cross-Cultural Research*, 46(2), pp.133-159.
- Mintzberg, H. (1994). *The rise and fall of strategic planning*. New York, NY: The Free Press.
- Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, 6(3), 257–272.
- Montes, C. (2003). *Results-Based Public Management in Bolivia* London: Overseas Development Institute.
- Morra, L.G. & Friedlander, A.C. (1999). *Case study evaluations*. Washington, DC: The World Bank.

- Motter, A., & Cox, A. (2010). *Making aid work: Towards better development results*. Practical guidance for parliamentarians on the role of parliaments in development effectiveness.
- Mowles, C. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation*, 20(2), 160-175.
- Moynihan, D. P. (2005). Goal-based learning and the future of performance Management. *Public Administration Review*, 65(2), 203–216.
- Moynihan, D. P. (2006). Managing for results in state government: Evaluating a decade reform. *Public Administration Review*, January/February 2006, 77–89.
- Mudacumura, G. M., & Haque, M. S. E. (2004). From development administration to new public management in postcolonial settings: Internal problems, external prescriptions. In M. Dekker (Ed.), *Handbook of Development Policy Studies*. New York, NY: Routledge.
- Nagel, S. S. (2002). *Public policy studies*. Huntington, NY: Nova.
- Naidoo, I. (2010). M&E in South Africa. Many purposes, multiple systems. In *From policies to results: Developing capacities for country M&E systems*, 301–322. New York, NY: UNICEF.
- Nakamura, A. O., & Warburton, W. P. (1998). Performance measurement in the public sector. *Canadian Business Economics*, Winter, Volume 6(2), 37–48.
- Neubert, S. (2000). *Social impact analysis of poverty alleviation programmes and projects: a contribution to the debate on the methodology of evaluation in development cooperation*. London, UK: Frank Cass.
- New Zealand. Ministry of Health. (2006). *A guide to developing public health programmes: a generic programme logic model*. Wellington, New Zealand: Ministry of Health.
- Newcomer, K. E. (1997). *Using performance measurement to improve public and non-profit programs*. San Francisco, CA: Jossey-Bass.
- Newcomer, K. E., & Downy, A. (1997–98). Performance-based management: What is it and how do we get there? *The Public Manager*, 26(4), 37–48.
- Norman, R. (2007). Managing outcomes while accounting for outputs: redefining public value in New Zealand's performance management system. *Public Performance & Management Review*, 30 (4), 536-4.
- Nunberg, B. (1992). *Managing the Civil Service: What LDCs can learn from developed country reform*. Policy Research Working Paper Series. Washington, DC: The World Bank.
- Nunns, H., Peace, R., & Witten, K. (2015). Evaluative reasoning in public-sector evaluation in Aotearoa New Zealand: How are we doing? *Evaluation Matters – He Take Tō Te Aromatawai (1) 2015*, 137–163.
- NZAID. (2006). *NZAID Activity cycle management guideline*. <https://nzaidtools.nzaid.govt.nz/tools>
- NZAID. (2006). *NZAID Logical Framework Approach Guideline*. <https://nzaidtools.nzaid.govt.nz/tools>
- Office of the Auditor General of Canada. (2000). *Implementing results-based management: Lessons from the literature*. [https://www.oag-bvg.gc.ca/domino/other.nsf/html/00rbm\\_e.html](https://www.oag-bvg.gc.ca/domino/other.nsf/html/00rbm_e.html)
- Office of the Auditor General of Canada. (2001). *Public service management reform: progress, setbacks, and challenges*. [https://www.oag-bvg.gc.ca/domino/other.nsf/html/01psm\\_e.html](https://www.oag-bvg.gc.ca/domino/other.nsf/html/01psm_e.html)



- Office of the Auditor General of Canada, Treasury Board of Canada Secretariat. (2003). *The managing for results self-assessment tool*. Ottawa, Canada, Treasury Board of Canada Secretariat.
- Ohara, S., & Asada, T. (Eds.). (2009). *Japanese project management: KPM – innovation, development and improvement*. Hackensack, NJ: World Scientific.
- Oliver, C. (2011). Critical realist grounded theory: A new approach for social work research. *British Journal of Social Work*, 42(2), 371-387.
- Organisation for Economic Co-operation and Development. (1997). *In search of results: Performance management practices*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2001). *The DAC guidelines. Poverty reduction*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2002). *The DAC guidelines: Integrating the Rio conventions into development co-operation*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2003). *Harmonizing donor practices for effective aid delivery. Good Practice Papers DAC Guidelines and Reference Series*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2005a). *Paris declaration of aid effectiveness*. Paris, France: OECD.  
<https://www.oecd.org/dataoecd/11/41/34428351.pdf>
- Organisation for Economic Co-operation and Development (2005b). *Managing for development results principles in action: Sourcebook on emerging good practice*. Working Party on Aid Effectiveness and Donor Practices Joint Venture on Managing for Development Results. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2006a). *Emerging good practice in managing for development results* (1<sup>st</sup> issue ed.) Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2006b). *Integrating human rights into development: Donor approaches, experiences and challenges*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2007). *Aid effectiveness: 2006 survey on monitoring the Paris Declaration*. Country chapters. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD (2008a). *Emerging good practice in managing for development results: Source book 3rd edition*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2008b). *Round table 4: Managing for development results*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2008c). *2008 survey on monitoring the Paris Declaration: Effective aid by 2010? What it will take?* Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2009a). *Civil society and aid effectiveness: findings, recommendations and good practice*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009b). *Implementing the Accra Agenda for Action "Beginning Now" commitments – updated compendium*. Paris, France: OECD.

- Organisation for Economic Co-operation and Development OECD. (2009c). *International good practice principles for country-led division of labour and complementarity. Working party on aid effectiveness*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009d). *Investing in development: A common cause in a changing world*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development OECD. (2009e). *Managing aid: practices of DAC member countries*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009f). *Making aid more effective through the strengthening and use of national systems*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009g). *Measuring Government. Activity*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009h). *Mutual accountability*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2009i). *Policy brief: Managing for development results*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2010a). *Assessing progress: 2011 survey on monitoring the Paris Declaration: Key facts*. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2010b). *Glossary of key terms in evaluation and results-based management*. DAC Working Party on Aid Evaluation. Paris, France: OECD.
- Organisation for Economic Co-operation and Development. (2010c). Country systems, and why we need to use them. Chapter Three in *Development Co-operation Report*. Paris, France: OECD. <https://doi.org/10.1787/dcr-2010-6-en>
- Organisation for Economic Co-operation and Development. (2010d). *Quality standards for development evaluation*. Paris, France: OECD.
- Organisation for Economic Development and Cooperation. (2011). *The high-level fora on aid effectiveness: A history*. The Fourth High Level Forum: Success factors. Paris, France: OECD. <https://www.oecd.org/document>
- Organisation for Economic Co-operation and Development. (2017). *How's Life? 2017: Measuring well-being*. Paris, France: OECD. [https://doi.org/10.1787/how\\_life-2017-en](https://doi.org/10.1787/how_life-2017-en)
- Organisation for Economic Co-operation and Development. (2019). *Learning from results-based management (RBM) evaluations and reviews*. Discussion paper for the OECD/ DAC Results Community Workshop on 29-30 October 2018, Paris, France.
- Organisation for Economic Co-operation and Development ECD & The World Bank. (2005). *Harmonisation, alignment, results: Progress report on aid effectiveness*. Washington, DC: The World Bank.
- Osei-Kojo, A., Bawole, J. N., & Sakyi, E. K. (2019). The opportunities and constraints to collaboration in public sector management. *Public Organization Review*.
- O'Sullivan, R. (2010). Collaborative evaluation. Creating environments to engage stakeholders in evaluation. In *From policies to results: Developing capacities for country M&E systems*, (pp. 289–300). New York, NY: UNICEF.
- Owen, J. M., & Rogers, P. J. (1999). *Program evaluation: forms and approaches* (International ed.). London, UK: Sage.

- Packer, M. J. (2011). *The science of qualitative research* (2nd ed.): Cambridge, MA: University Press.
- Page C, & Ayres R. Policy Logic: Creating policy and evaluation capital in your organisation. *Evaluation Journal of Australasia*. 2018;18(1):45-63.  
<https://doi:10.1177/1035719X17753960>
- Pal, L. A. & Teplova, T., (2003). *Rubik's Cube? Aligning organizational culture, performance measurement and horizontal management*. Ottawa, Canada: Carleton University.
- Papua New Guinea Consolidated Legislation. (1975). *Constitution of the Independent State of Papua New Guinea*.  
[https://www.paclii.org/pg/legis/consol\\_act/cotisopng534/](https://www.paclii.org/pg/legis/consol_act/cotisopng534/)
- Patrizi, P. A. (2010). Strategy evaluation: Emerging processes and methods. *New Directions for Evaluation*, 2010(128), 87-102.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: The Guilford Press.
- Patton, M. Q. (2016). What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-Term commitment, integrity, and sensitivity in implementing evaluation models. *American Journal of Evaluation*, 37(2), 250-265.
- Patton, M. Q & Patrizi, P. A. (2010). Strategy as the focus for evaluation. *New Directions for Evaluation*, 2010, 5-28.
- Pawson, R. (2006). *Evidence-based policy. A realist perspective*. London, UK: Sage.
- Pawson, R. (2008). Invisible mechanisms. *Evaluation Journal of Australia*, 8(2), 3–13.
- Pawson, R. (2013). *The science of evaluation: A realist manifesto*. London, UK: Sage.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: an introduction. ESRC Research Methods Programme University of Manchester, RMP Methods Paper 2/2004*.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*: London, UK: Sage.
- Pawson, R., & Tilley, N. (2001). Realistic evaluation bloodlines. *American Journal of Evaluation*, 22(3), 317–324.
- Pazakavabwa, A. & Steyn, G. M. (2014). Results-based management in the public sector of developing countries: What should be considered? *Mediterranean Journal of Social Sciences*, 5(20), 245.
- Perrin, B. (1998). Effective use and misuse of performance measurement. *American Journal of Evaluation*, 19(3), 367.
- Perrin, B. (1999). Performance measurement: Does the reality match the rhetoric? A rejoinder to Bernstein and Winston. *American Journal of Evaluation*, 20(1), 101.
- Perrin, B. (2002). *Implementing the vision: Addressing challenges to results-focussed management and budgeting*. Paris, France: OECD.
- Perrin, B. (2006). *Moving from outputs to outcomes: Practical advice from governments around the world. World Bank managing for performance and results series*. Washington DC. The World Bank.
- Peters, B. G., & Pierre, J. (1998). Governance without government? Rethinking public administration. *Journal of Public Administration Research & Theory (Transaction)*, 8(2), 223.

- Peters, G., & Savoie, D. J. (Eds.). (2000). *Governance in the twenty-first century: revitalizing the public service*. Montreal, Canada: Canadian Centre for Management Development.
- Phillips, L., & Ilcan, S. (2004). Capacity-building: The neoliberal governance of development. *Canadian Journal of Development Studies*, XXV(3).
- Picciotto. (2009). Evaluating development: Is the country the right unit of account? In *Country-led M&E systems: Better evidence, better policies, better development results*, (pp. 32–55). New York, NY: UNICEF.
- Picciotto, R. (2002). The logic of mainstreaming: A development evaluation Perspective. *Evaluation*, 8(3), 322–339.
- Picciotto, R. (1999). *Towards an economics of evaluation* (Vol. 5). London, UK: Sage.
- Pieterse, J. N. (2000). After post-development. *Third World Quarterly*, 21(2), 175-191.
- Pieterse, J. N. (2010). *Development theory. Deconstructions/reconstructions* (2<sup>nd</sup> ed.). London, UK: Sage.
- Pitman, G. K., Feinstein, O. N., & Ingram, G. K. (Eds.). (2005). *Evaluating development effectiveness* (Vol. 7). New Brunswick, NJ: Transaction.
- Plantz, J., Greenway, M., & Hendricks, M. (1997). Outcome measurement: showing results in the Non-profit Sector. *New Directions for Evaluation*, Fall (75), 15–30.
- Poate, D. (1997). *Measuring & managing results: Lessons for development cooperation*. New York, NY: OECD.
- Poister, T. H. (2003). *Measuring performance in public and non-profit organizations*. San Francisco, CA: Jossey-Bass John Wiley.
- Pollitt, C. (1998). Managerialism revisited, In B.G. Peters. and D. J. Savoie (Eds.) *Taking stock: Assessing public sector reform*. Montreal, Canada: Canadian Centre for Management Development.
- Popper. (1963). *Conjectures and refutations*: London, UK: Routledge.
- Porter, S., & Goldman, I. (2013). A growing demand for M&E in Africa. *African Evaluation Journal*, 1. <https://doi:10.4102/aej.v1i1.25>
- Power, D. (2011). What is Henry Mintzberg's view of strategic planning? 2011 (August 8, 2011).
- Powers, L. C. (2009). *A framework for evaluating the effectiveness of performance measurement systems*. *Real World Systems Research Series 2009*, 1. <https://ssrn.com/abstract=1371158>
- Prashnig, B. (1998). *The power of diversity: new ways of learning and teaching*. Auckland, New Zealand: David Bateman.
- Preston, P. W. (1996). *Development theory: an introduction*. Oxford, UK: Blackwell.
- Pulotu-Endermann, F. K. (2001). *Fonofale Model of Health*. Unpublished manuscript. Wellington, New Zealand.
- Quesnel, J. S. (2009). The strategic intent. Understanding strategic intent is the key to country-led monitoring and evaluation systems. In Segone, M. (Ed) *Country-led M&E systems: Better evidence, better policies, better development results*, (pp. 56–76). New York, NY: UNICEF.
- Rai, N., Smith, B., & Brooks, N. (2019). *Assessing adaptation results. Aligning national M&E systems and global results frameworks. Issues paper*. International Institute for Environment and Development. <https://pubs.iied.org/10198IIED/>

- Rainero, C., & Brescia, V. (2018). The participatory budgeting towards a new governance and accountability. *International Journal of Management Sciences and Business Research*, 7(2).
- Reeler, D. (2007). *A three-fold theory of social change and implications for practice, planning, monitoring and evaluation*. Cape Town, South Africa: The Community Development Resource Association Centre (CDRA).
- Reichert, J. (2010). Abduction: The logic of discovery of grounded theory. *Forum: Qualitative Social Research Sozialforschung*, 11(1).
- Renger, R. (2015). System evaluation theory (SET): A practical framework for evaluators to meet the challenges of system evaluation. *Evaluation Journal of Australasia*, 15(4), 16-28. <https://doi.org/10.1177/1035719X1501500403>
- Renger, R., Atkinson, L., Renger, J., Renger, J., & Hart, G. (2019). The connection between logic models and systems thinking concepts. *Evaluation Journal of Australasia*, 19(2), 79-87. <https://doi.org/10.1177/1035719x19853660>
- Renzio, P. D. & Lakin, J. (2019) *Reframing public finance: Promoting justice, democracy, and human rights in government budgets*. International Budget Partnership. <https://www.internationalbudget.org/publications/reframing-public-finance-promoting-justice-democracy-human-rights-in-government-budgets/>
- Reynolds, M. (2008). Response to paper “systems thinking” by D. Cabrera et al., Systems thinking from a critical systems perspective. *Discussion / Evaluation and Program Planning*, 31, 311–334.
- Reichert, E. (2006). *Understanding human rights: An exercise book*. Thousand Oaks, CA: Sage.
- Rist, G. (1997). *The history of development: From western origins to global faith*. London, UK: Zed.
- Rist, R. C. (2006). *Building a results-based management system to measure development results*. Washington, DC: The World Bank.
- Rist, R. C. (2009). Country-led M&E systems. Better evidence, better policies better development results. Preface. In *Country-led M&E systems: Better evidence, better policies, better development results* (pp. 32–55). New York: UNICEF.
- Robinson, D. (1999). The development management task and reform of ‘public’ social services. *Development in Practice*, 9(1&2), 78–85.
- Robinson, M., & Brumby, J. (2005). *Does performance budgeting work? An analytical review of the empirical literature*. Washington: International Monetary Fund.
- Rodril, D. (2011). The Future of Convergence. NBER Working Paper No. 17400. Cambridge MA: Harvard University.
- Rogers, P. J. (ed.) (2000). *Program theory in evaluation: challenges and opportunities*. San Francisco, CA: Jossey-Bass.
- Rogers, P. J. (2008a). Response to paper “systems thinking” by D. Cabrera et al.: Is it systems thinking or just good practice in evaluation? *Evaluation and Program Planning*, 31(3), 325–326. <https://doi.org/10.1016/j.evalprogplan.2008.04.005>
- Rogers, P. J. (2008b). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48.
- Rogers, P. J., Hasci, T. A., Petrosino, A., & Huebner, T. A. (Eds.). (2000). *Program theory in evaluation: challenges and opportunities*. San Francisco, CA: Jossey-Bass.
- Rogers, P. P., Jalal, K. F., & Boyd, J. A. (2008). *An introduction to sustainable development*. London Sterling, VA: Earthscan.

- Rolfe, S. (2019). Combining theories of change and realist evaluation in practice: Lessons from a research on evaluation study. *Evaluation*, 25(3), 294-316. <https://doi.org/10.1177/1356389019835229>
- Rossi, P. H., Lipsey, M. W. & Freeman, H. E. (2004). *Evaluation: a systematic approach* (7th ed.). Thousand Oaks, CA: Sage.
- Rostow, W.W. (1960). *The stages of economic growth*. (1<sup>st</sup> ed.). Cambridge, UK: Cambridge University Press.
- Rouse, A. P. B. (1997). *A methodological framework of performance measurement with applications using data envelopment analysis*. (PhD thesis). The University of Auckland, Auckland.
- Ryan B. (2003). 'Death by evaluation'? Reflections on monitoring and evaluation in Australia and New Zealand. *Evaluation Journal of Australasia*. 2003;3(1):6-16. <https://doi.org/10.1177/1035719X0300300103>
- Ryan, B. (2011). Public management in difficult economic times. *Policy Quarterly*, 7(3).
- Sachs. (2012). From millennium development goals to sustainable development goals. *The Lancet*, Vol. 379, Issue 9832, 2206– 2211. -
- Samoa Ministry of Finance. (2008). *Strategy for the development of Samoa 2008–2012. Ensuring sustainable economic and social progress*. Apia, Samoa
- Sause, L. L. (2008). *Policy advisory capability in Papua New Guinea's central government: evaluation, implications and lessons*. (Doctor of Philosophy in Public Policy). Wellington: Victoria University of Wellington.
- Schick, A. (1996). *The spirit of reform: Managing the New Zealand state sector in a time of change*. Wellington, New Zealand: State Services Commission.
- Schick, A. (1998). Why Most Developing Countries Should Not Try New Zealand's Reforms. *World Bank Research Observer*, 13(1), 123-131.
- Schneider, S. C. and J.-L. Barsoux. (1997). *Managing across cultures*. London, UK: Prentice Hall.
- Schon, D. (1963). *Invention and the Evolution of Ideas*: London, UK: Associated.
- Schon, D. A (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic Books.
- Schouten, P. (2009). Theory Talk #34: James Ferguson on Modernity, Development, and Reading Foucault in Lesotho. *Theory Talks*. <https://www.theory-talks.org/2009/11/theory-talk-34.html>
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In *Handbook of qualitative research* (pp. 118-137).
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118–137). Thousand Oaks, CA: Sage.
- Schwandt, T. A. (2005). The centrality of practice to evaluation. *American Journal of Evaluation*, 26, 95–105. <https://doi.org/10.1177/1098214004273184>
- Schwandt, T.A. (2015). *Evaluation foundations revisited: Cultivating a life of the mind for practice*. Stanford, CA: Stanford University Press.
- Schwandt, T., Ofir, Z., D'Errico, S., El-Saddik, K., & Lucks, D. (2016). *Realising the SDGs by reflecting on the way(s) we reason, plan and act the importance of evaluative thinking*. International Institute for Environment and Development. <https://doi.org/10.13140/RG.2.2.11995.59683>

- Scott, G. C. (2001). *Public management in New Zealand: Lessons and challenges*. Wellington, New Zealand: New Zealand Business Roundtable.
- Scott, K. W. (2004). Relating categories in grounded theory analysis: Using a conditional relationship guide and reflective coding matrix. *The Qualitative Report*, 9(1), 113–126.
- Scott, L., Joubert, A., & Anyogu, A. (2005). *Working Paper no. 2: Setting National Goals, Monitoring and Evaluation: a comparison of approaches across selected countries*. Cape Town, South Africa: Department of Statistical Sciences, University of Cape Town.
- Scriven, M. (1991). Pros and cons about goal-free evaluation. *American Journal of Evaluation*, 12(55), 55–62.
- Scriven, M. (1997). Empowerment evaluation examined. *Evaluation Practice*, 18(1), 165-175. <https://doi:10.1177/109821409701800115>
- Scriven, M. (2007). The logic of evaluation. In H.V. Hansen, et. al. (Eds.), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1-16). Windsor, ON: OSSA.
- Segone, M. (Ed.). (2009). *Country-led monitoring and evaluation systems: Better evidence, better policies, better development results*. New York, NY: UNICEF
- Sen, A. (1999). *Development as freedom*. New York, NY. Anchor Books.
- Sen, A. (2002). Response to Commentaries. *Studies in Comparative International Development*, 37(2), 78. <https://doi:10.1007/BF02686264>
- Serrat, O. (2008). Outcome mapping. *Knowledge Solutions*, November 2008(17). [https://www.researchgate.net/publication/254583227\\_Outcome\\_Mapping](https://www.researchgate.net/publication/254583227_Outcome_Mapping)
- Shadish, W. R., & Luellen, J. K. (2004). Donald Campbell: The accidental evaluator. In M. C. Alkin (Ed.), *Evaluation roots: tracing theorists' views and influences* (pp 80–87). London, UK: Sage.
- Shaw, I., Greene, J. C., & Mark, M. M. (Eds.). (2006). *The SAGE handbook of evaluation: policies, programs and practices*. London, UK: Sage.
- Shutt, C. (2016). *Towards an alternative development paradigm?* .Rapport 2016:07 till Expertgruppen för biståndsanalys (EBA). London, UK: Institute of Development Studies.
- Smith, M. K. (2001). Chris Argyris: theories of action, double-loop learning and organizational learning. *The encyclopaedia of informal education*. <https://www.infed.org/thinkers/argyris.htm>
- Smith, M. K. (2001). Donald Schon: learning, reflection and change. *The encyclopaedia of informal education*. <https://www.infed.org/thinkers/et-schon.htm>.
- Smith, N. L., & Brandon, P. R. (eds.). (2008). *Fundamental issues in evaluation*. New York, NY: Guilford Press.
- Social Investment Agency (SIA). (2018). *Are we making a difference in the lives of New Zealanders – how will we know? A wellbeing measurement approach for investing for social wellbeing in New Zealand*. SIA Working Paper. Wellington, New Zealand: New Zealand Government.
- Social Security Office, T. (n.d.). *Results-Based Management (RBM) System*. Bangkok, Thailand: Social Security Office.
- Spiller C., Barclay-Kerr, H., & Panoho, J. (2016). *Wayfinding leadership: Ground-breaking wisdom for developing leaders*. Wellington, New Zealand: Huia.
- Stake, R. E. (2004). *Standards-based & responsive evaluation*. Thousand Oaks, CA: Sage.

- State Sector Act. (1988). *New Zealand: State Services Commission*. Wellington, New Zealand. <https://www.legislation.govt.nz/act/public/1988/0020/latest/whole.html>
- State Services Commission. (1998). A better focus on outcomes through SRA networks. *Occasional Paper, No. 3*. Wellington, New Zealand.
- State Services Commission. (1999). Assessing departments' capability to contribute to strategic priorities. *Occasional Paper, No. 16*. Wellington, New Zealand.
- State Service Commission (2003a). *Doing the right things and doing them right: Improving evaluative activity in the New Zealand state sector*. Wellington, New Zealand. <https://www.scc.govt.nz/display/document.asp?DocID=3507>
- State Services Commission. (2003b). *Guidance on outcomes-focused management – building block 3: Intervention Logic (version 2.1)*. Wellington, New Zealand.
- State Services Commission. (2004). *Getting better at Managing for Shared Outcomes – a resource for agency leaders*. Wellington, New Zealand. [https://www.ssc.govt.nz/display/document.asp?docid=4126&pageno=1#P5\\_0](https://www.ssc.govt.nz/display/document.asp?docid=4126&pageno=1#P5_0)
- New Zealand State Services Commission. (2006a). *State of Development Goals: Report 2006*. Wellington, New Zealand.
- State Services Commission. (2006b). *State Services Commission Statement of Intent 2006*. Wellington, New Zealand.
- State Services Commission. (2007). *Performance information measures and standards in the SOI and annual report*. Wellington, New Zealand. <https://www.ssc.govt.nz/performance-info-measures>
- State Services. (2008a). *Performance measurement: Advice and examples on how to develop effective frameworks*. Wellington, New Zealand.
- States Services Commission. (2008b). *Performance measurement in support of decision-making: a summary*. Wellington, New Zealand.
- State Services Commission. (2016a). *Better public services - results for New Zealanders*. Wellington, New Zealand. <https://www.ssc.govt.nz/bps-results-for-nz>
- Statistics New Zealand. (2014). *New Zealand in Profile, 2014*. Wellington, New Zealand.
- Steer, L., Wathne, C., & Driscoll, R. (2009). *Mutual accountability at the country level. a concept and emerging good practice paper*. London: ODI.
- Steering Group for the Managing for Outcomes Roll-out 2004/05. (2003). *Learning from evaluative activity: Enhancing performance through outcome-focused management*. Wellington, New Zealand: State Services Commission.
- Steering Group for the Managing for Outcomes Roll-out 2003/4. (2002). *Managing for Outcomes. Guidance for Departments*. Wellington, New Zealand: State Services Commission.
- Steering Group for the Managing for Outcomes Roll-out 2004/5. (2003). *Managing for Outcomes Guidance*. Wellington, New Zealand: State Services Commission.
- Stern, E. D. (1999). *Editorial (5). Why Parliament should take evaluation seriously. The Evaluator*. London, UK: Sage.
- Stern, E. D. (2003). *The rationale & challenges for evaluation of public policies*. Paper presented at the Fifth European Conference on Evaluation of Structural Funds, Budapest.
- Stern, E. D. (2004). Philosophies and types of evaluation research. In P. Descy, & M. Tessaring (Eds.), *Cedefop. Foundations of Evaluation and Impact Research*. Luxembourg. Office of Official Publications of the European Communities.



- Stern, E. D. (2008a). Editorial (Vol. 5). *Evaluation*. London, UK: Sage.
- Stern, E. D. (2008b). *Thematic study on the Paris Declaration, aid effectiveness and development effectiveness*. Koege, Denmark: OECD.  
<https://www.oecd.org/dac/evaluationnetwork>
- Stickl Haugen, J., & Chouinard, J. A. (2018). Transparent, translucent, opaque: Exploring the dimensions of power in culturally responsive evaluation contexts. *American Journal of Evaluation*, 40(3), 376–394.  
<https://doi.org/10.1177/1098214018796342>
- Stockmann, R., & Meyer, W. (2016). The future of evaluation: Global trends, new challenges and shared perspectives. *The Future of Evaluation*, 9–19.  
[https://doi.org/10.1057/9781137376374\\_1](https://doi.org/10.1057/9781137376374_1)
- Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7–98. <https://doi.org/10.1002/ev.3>
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications* (1st ed.). San Francisco, CN: Jossey-Bass.
- Sutcliffe, S., & Court, J. (2005). *Evidence-based policymaking. What is it? How does it work? What relevance for developing countries?* London, UK: Overseas Development Institute.
- Superu. (2017). *Making sense of evaluation: A handbook for everyone*. Wellington, New Zealand: Social Policy Evaluation and Research Unit (Superu).  
[https://thehub.swa.govt.nz/assets/documents/V2\\_Handbook\\_FINAL-enhanced.pdf](https://thehub.swa.govt.nz/assets/documents/V2_Handbook_FINAL-enhanced.pdf)
- Swiss, J. E. (2005). A framework for assessing Incentives in results-based management. *Public Administration Review*, 65(5), 592–602.
- Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social & behavioural research*. Thousand Oaks, CA: Sage.
- Tawhai, V. M. & Gray-Sharp, K. (2011). *'Always speaking': The Treaty of Waitangi and public policy*. Wellington, New Zealand: Huia.
- Te Ara: The Encyclopaedia of New Zealand. (2006). *Māori Peoples of New Zealand*. Auckland, New Zealand: David Bateman.
- The Alotau Accord. (2012). *Summary of O' Neill Government Priorities*. Government of Papua New Guinea.
- The Alotau Accord 2. (2017). *Summary of O' Neill Government Priorities*. Government of Papua New Guinea.
- The International Bank for Reconstruction. (2000). *Designing and building a results-based M&E system: A tool for public sector management*. Washington, DC: The World Bank.
- The Treasury & State Services Commission. (2005a). *Getting better at managing for outcomes: A tool to help organisations consider their progress in results-based management and identify development objectives*. Wellington, New Zealand: The Treasury.
- The Treasury & State Services Commission. (2005b). *Guidance for Crown Entities: Planning and managing for results*. Wellington, New Zealand: The Treasury.  
[https://www.ssc.govt.nz/upload/downloadable\\_files/planning-and-managing-for-results-](https://www.ssc.govt.nz/upload/downloadable_files/planning-and-managing-for-results-)
- The Treasury. (2019). *The wellbeing budget, 2019*. Government of New Zealand. Wellington, New Zealand.

- The World Bank. (2002). *Better measuring, monitoring, and Managing for Development Results*. Development Committee (Joint Ministerial Committee of the Boards of Governors of the World Bank and the International Monetary Fund on the Transfer of Resources to Developing Countries).  
[https://siteresources.worldbank.org/DEVCOMMINT/Documentation/90015418/D C2002-0019\(E\)-Results.pdf](https://siteresources.worldbank.org/DEVCOMMINT/Documentation/90015418/D C2002-0019(E)-Results.pdf)
- The World Bank (2005). *Equity and Development*. World Development Report. New York, NY: The World Bank and Oxford University Press.
- The World Bank (2007). *The country based PREM poverty reduction group. development model and scaling up. The Scaling Up Newsletter* (2). Washington, DC.
- The World Bank. (2008a). Multilateral Development Banks' Common Performance Assessment System: COMPAS 2008 Report. Washington, DC.
- The World Bank. (2008b). *Using knowledge to improve development effectiveness: an evaluation of the World Bank economic and sector work and technical assistance, 2000–2006*. Washington, DC: Independent Evaluation Group, World Bank.
- Thomas, D. (2002). Evaluating the cultural appropriateness of service delivery in multi-ethnic communities. *Evaluation Journal of Australasia*, 22(new series), 50–56.
- Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry*, 17(6), 511-521.
- Thomas, T. & Duncan, L., (Eds). (2010). *Papua New Guinea's development performance, 2005- 2008*. Port Moresby: National Research Institute.
- Thomas, V. (2009). *Evaluation to help improve development results*. Informal paper. Washington, DC: The World Bank.
- Tilley, N. (2000). *Realistic evaluation: An overview*. Paper presented at the Founding Conference of the Danish Evaluation Society. <https://fm8-10042.nt.uni2dk/pdf/Nick%20Tilley.pdf>
- Trauer, T. (ed.) (2010). *Outcome measurement in mental health: theory and practice*. Cambridge, UK: Cambridge University Press.
- Treasury and the State Services Commission. (2005). *Getting better at managing for outcomes: a tool to help organisations consider their progress in results-based management and identify development objectives*. Wellington, New Zealand.
- Treasury Board of Canada Secretariat. (2001). *Guide for the development of results-based management and accountability frameworks*. Ottawa, Canada.
- Treasury H M. (2003). *The green book: appraisal and evaluation in central government*. London, UK. <https://www.hm-treasury.gov.uk>.
- Tucker, V. (1999). The myth of development: A critique of eurocentric discourse. In R. Munck and D. O'Hearn (Eds.), *Critical development theory: Contributions to the new paradigm* (pp. 1-26). London, UK: Zed Books.
- Tumusiime-Mutebile, E. (2002). *Managing for Development Results*. Paper presented at the OECD-DAC Development Partnership Forum. Paris, France: OECD.
- United Nations Development Programme Evaluation Office. (2002). *Handbook on monitoring and evaluating for results*. New York, NY: UNDP.
- United Nations Development Programme. (2007). *Evaluation of results-based management at UNDP*. New York, NY: United Nations Development Programme.

- United Nations Evaluation Group (UNEG). (2007a). *Core competencies for evaluators of the UN system*. New York, NY: United Nations Evaluation Group. <https://www.uneval.org/document/detail/1408>
- United Nations Evaluation Group (UNEG). (2007b). *The role of evaluation in results-based management. Reference document, UNEG/REF(2007)1*. New York, NY: UNEG.
- United Nations Development Programme. (2010a). *Millennium development goals report 2010*. New York, NY: UNDP
- United Nations Development Programme. (2010). *What will it take to achieve the Millennium Development Goals? – An international assessment*. New York, NY: UNDP.
- United Nations Development Programme. (2014). *Papua New Guinea national human development report 2014: from wealth to wellbeing: translating resource revenue into sustainable human development*. Port Moresby, Papua New Guinea: UNDP.
- United Nations Development Programme. (2015a). Managing the transition from the millennium development goals to the sustainable development goals: what it will take? *Report of the Secretary-General 2015 session 21 July 2014 – 22 July 2015, Agenda item 5(c)*. New York, NY: UNDP. <https://pdfs.semanticscholar>
- United Nations Development Programme. (2015b). *Transforming our world: the 2030 agenda for sustainable development*. New York, NY: UNDP. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>
- United Nations Development Programme. (2016). *Pacific risk resilience programme: a partnership approach to risk governance in the Pacific*. Sustainable Development Goals, partnerships platform. <https://sustainabledevelopment.un.org/partnership/?p=7381#updates>
- United States General Accounting Office (1997). *The Government Performance and Results Act: 1997 Government-wide implementation will be uneven*. GAO/GGD-97-109. Washington, DC.
- United States Office of Auditor General. (2005). *Performance measurement and evaluation. Definitions and relationships*. Washington, DC.
- United Nations Research Institute for Social Development (UNRISD). (2015). *Revisiting sustainable development*. Geneva, Switzerland: UNRISD.
- United Nations World Commission on Environment and Development. (1987). *Our common future: report of the World Commission on Environment and Development*. New York, NY: UN.
- Vaessen, J., & Leeuw, F. L. (Eds.). (2010). *Mind the gap: perspectives on policy evaluation and the social sciences*. New Brunswick, NJ: Transaction.
- Vail, J. (2007). Community-based development in Tari: Present and prospects. In N. Haley, & R. May (Eds.), *Conflict and Resource Development: in the Southern Highlands of Papua New Guinea* (pp. 107–122). Canberra, Australia: ANU E Press.
- Van Thiel, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance and Management Review*, 25(3), 267–281.
- Vasilachis de Gialdino, Irene (2011). Ontological and epistemological foundations of qualitative research. *Forum: Qualitative Social Research*, 10(2), Art. 30, <https://nbn-resolving.de/urn:nbn:de:0114-fqs0902307>

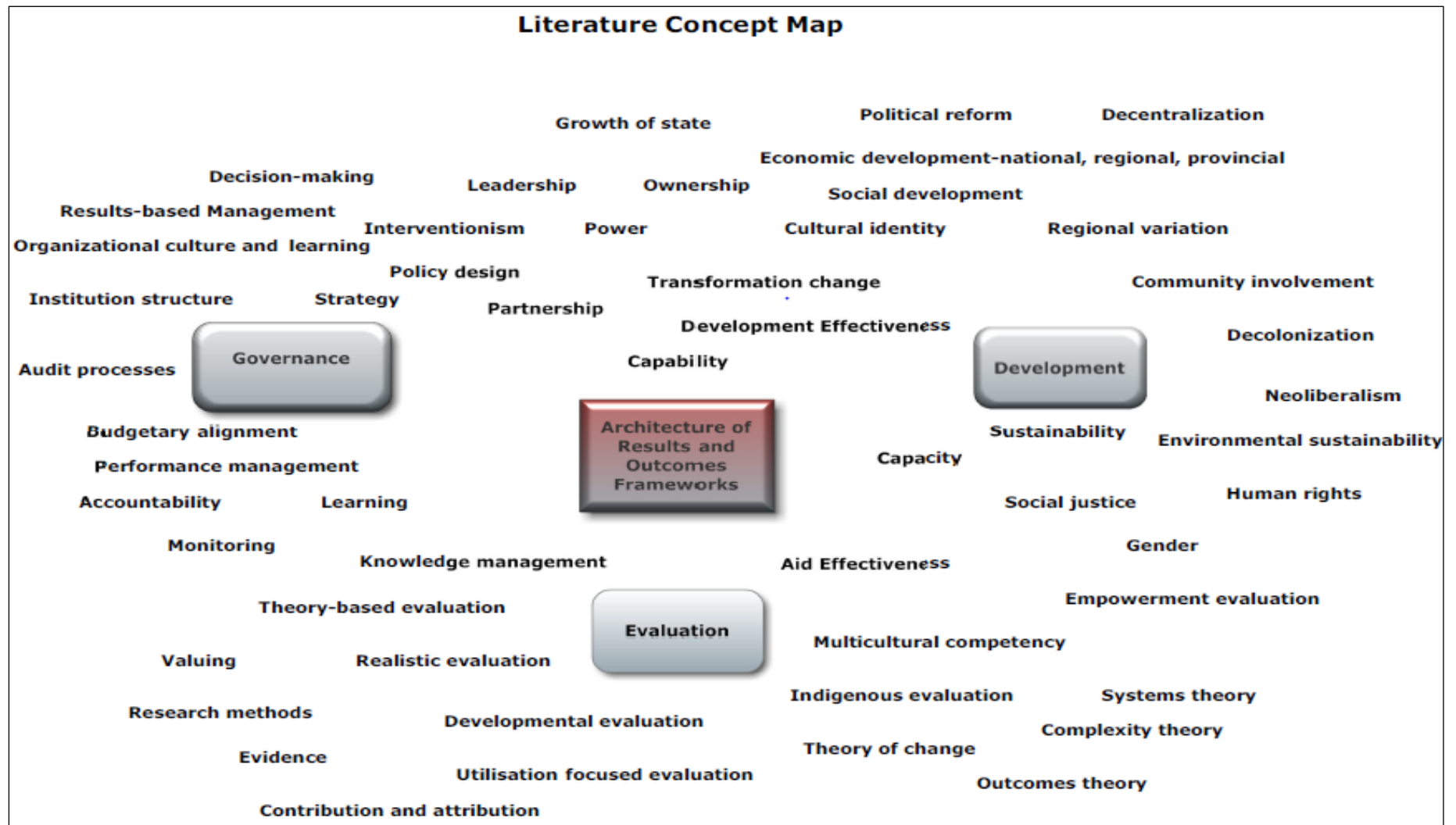
- Vedung, E. (1997) *Public policy and program evaluation*. New Brunswick, NJ: Transaction Publishers.
- Venelasen, L., & Miller, V. (2002). *A new weave of power, people, and politics: the action guide for advocacy and citizen participation*. Bourton-on-Dunsmore. 10.3362/9781780444208.
- Vidal, C. (2008) Wat is een wereldbeeld? (What is a worldview?). In H Van Belle, & J. Van der Veken (Eds.), *Nieuwheid denken. De wetenschappen en het creatieve aspect van de werkelijkheid*. Acco, Leuven, Belgium: In press.
- Wagner, R.B. 1989. *Accountability in education: a philosophical inquiry*. London, UK: Routledge.
- Ward, A. (1991). Interpreting the Treaty of Waitangi: the Māori resurgence and race relations in New Zealand. *The Contemporary Pacific*, 3(1), 85-113. <https://www.jstor.org/stable/23701489>
- Watts, M. (2006). Culture, development and global neo-liberalism. Chapter 2 In *Culture and Development in a Globalizing World*. London, UK: Routledge. <https://doi.org/10.4324/9780203641019>
- Watt, T. (2006). *Central agency reform in Papua New Guinea*. Case program. The Australia and New Zealand School of Government. Wellington, New Zealand.
- Webster, K., & Cheyne, C. (2017). Creating treaty-based local governance in New Zealand: Māori and Pākehā views. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 12(2), 146–164. <https://doi.org/10.1080/117783X.2017.1345766>
- Weiss, C. H. (1972). *Evaluation research: methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.
- Weiss, C. H. (Ed.) (1977). *Using social research in public policy making*. Lexington, MS: Lexington Books.
- Weiss, C. H. (1998). *Evaluation: methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Weiss, C.H. (2000). 'Which links in which theories shall we evaluate?' *New Directions for Evaluation* 87(Fall), pp. 35–45. <https://doi.org/10.1002/ev.1180>
- Weiss, C. H., & Barton, A. H. (Eds.). (1980). *Making bureaucracies work*. Beverly Hills, CN: Sage.
- Wheen, N. R., & Haywood, J. (2012). *Treaty of Waitangi settlements*. Wellington, New Zealand: Bridget Williams Books.
- White, E. & Rodriguez, R. (2004). Results-oriented country programming: applying the principles of managing for results and emerging practices and lessons. *Managing for Development Results in Action: Sourcebook on Emerging Good Practices*. Washington, DC: World Bank.
- White, H. (2009). *The road to nowhere? Result-based management in international cooperation: Current issues, food for thought, background information*. [http://www.euforic.org/detail\\_page.phtml?lang=en&page=resource\\_briefing\\_evaluat\\_e](http://www.euforic.org/detail_page.phtml?lang=en&page=resource_briefing_evaluat_e)
- White, S. (2003). Freedom, Reciprocity, and the Citizen's Stake. In, K Dowding, J. De Wispelaere, & S. White, S. (Eds.), *The Ethics of Stake-holding*. London, UK: Palgrave Macmillan.
- Whitfield, L. (2009). Ownership and the donor-recipient relationship. *DIIS Working Paper*, 2009(18). <https://hdl.handle.net/10419/44699>
- Wholey, J. S. (1999). Performance-based management: Responding to the challenges. *Public Productivity and Management Review*, 22(3), 288–307.

- Wholey, J. S., & Newcomer, K. E. (1997). Clarifying goals, reporting results. *New Directions for Evaluation, Fall (75)*, 91–98.
- Williams, B (2005). Systems thinking. In, S. Mathison (Ed.) *Encyclopaedia of Evaluation*. Thousand Oaks, CA: Sage.
- Williams, B. (2015). Prosaic or profound? The adoption of systems ideas by impact Evaluation. *IDS Bulletin, 46*(1), pp. 7-16. <https://doi:10.1111/1759-5436.12117>
- Williams, B, & Hummelbrunner, R. (2011). *Systems concepts in action: A practitioner's toolkit*. Stanford, CA: Stanford University Press.
- Williams, B., & Iman, I. (Eds.). (2006). *Systems concepts in evaluation: an expert anthology*. Point Reyes, CA: EdgePress.
- Williams, P. & Chrisman, L. Eeds.). (1993). *Colonial discourse and post-colonial theory*. New York, NY: Harvester Wheatsheaf.
- Williamson, T. (2003). *Targets and results in public sector management: Uganda case study* (Working Paper 205). London, UK: Overseas Development Institute.
- Winston, J. A. (1999). Performance indicators-promises unmet: A response to Perrin. *American Journal of Evaluation, 20*(1), 95–100.
- Wollmann, H. E., Elgar, E. E., Pollit, C., & Bouckaert, G.(Eds.) (2003). Evaluating public management reforms: an international perspective in evaluation *In Public-Sector Reform. Concepts and Practice in International Perspective*. Berlin, Germany: Social Science Institute, Humboldt University of Berlin.
- Wood, B., Kabell, D., Sagasti, F., & Muwanga, N. (2008). *Evaluation of the Implementation of the Paris Declaration. Phase one. Synthesis report*. Copenhagen, Denmark: Ministry of Foreign Affairs of Denmark.
- Wren, J. (2007). How to monitor for public health outcomes: Guidelines for Developing a Monitoring Framework. *Public Health Intelligence: Occasional Bulletin No. 44*. Wellington, New Zealand: Ministry of Health.
- Wright, E. A. (2002). Re-engineering of interlocking log frames – An irrigation project design case. *Physics and Chemistry of the Earth, 27*(11–22), 1005–1010. <https://www.scopus.com/inward/record.url?eid=2-s2.0-0036422012&par>
- Wuisman, J. (2005). The logic of scientific discovery in critical realist social scientific research. *Journal of Critical Realism, 4*(2), 66-394.
- Wyatt Knowlton, L., & Phillips, C. C. (2009). *The logic model guidebook: Better strategies for great results*. Thousand Oaks, CA: Sage.
- Yin, R. K. (2009). *Case study research: design and methods* (4th ed.,vol.5). Los Angeles, CA: Sage.
- Zanini, G. (2001). *India, the challenges of development: a country assistance evaluation*. Washington, DC: The World Bank.

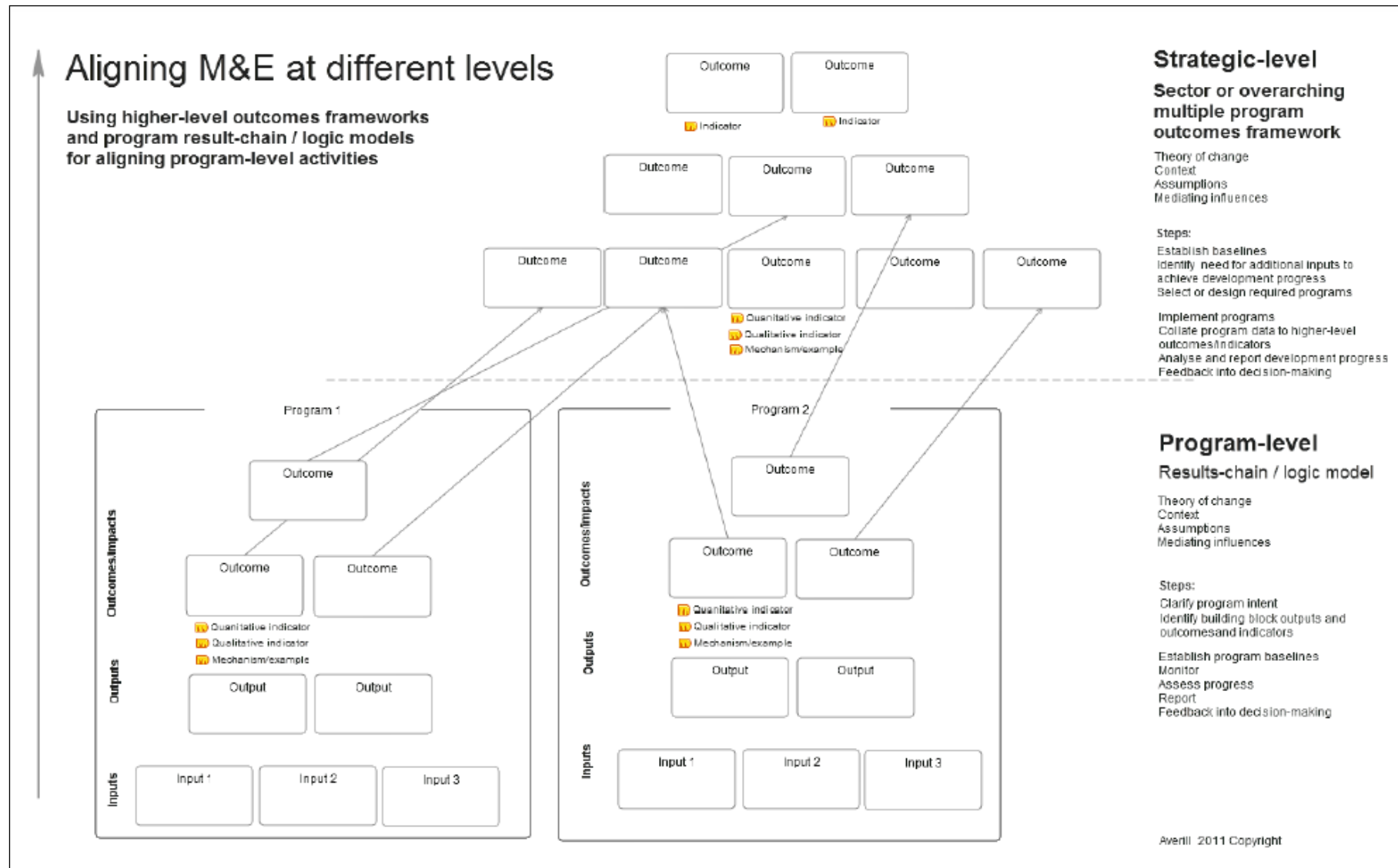
# Appendices

Appendix A: Literature review concept map (2011) .....	288
Appendix B: Aligning concepts at different levels .....	289
Appendix C: Ethics application .....	290
Appendix D: Data collection tools .....	293
Appendix E: Sample of qualitative findings analysis .....	305
Appendix F: Qualitative research findings thematic summary .....	325
Appendix G: Full summary table of identified concepts from research ....	333
Appendix H: List of organisations who participated in this research .....	342

## Appendix A: Literature review concept map (2011)



## Appendix B: Aligning concepts at different levels





## **Appendix C: Ethics application**

According to Massey University Ethical Guidelines, this research project met the criteria for Low-Risk Notification, based on the Screening Document. The Massey University Code of Ethical Conduct for Research, Teaching and Evaluations Involving Participants were followed. The ethical application was received on 13 December 2010.

The Code of Ethics and Guidelines for Ethical Conduct of Evaluations of the Australasian Evaluation Society was adhered to during this research project. The application and the low-risk notification form were peer reviewed and discussed within the Development Studies Department. It was agreed that this application for PhD research met the low-risk notification criteria. One issue was raised regarding informed and voluntary consent from managers and employees. Particular care was taken during the research to safeguard the confidentiality and emotional safety of research participants.

The following information was provided as part of the ethics application.

### **Summary of research project**

This PhD research will examine the principles underpinning evidence-based country and sector results frameworks. The primary research will focus on the architecture of these frameworks in New Zealand and a Pacific country international case studies). The principles underpinning country and sector results frameworks in different settings will be identified. The research will contribute to the knowledge base in development, policy, aid, management, monitoring and evaluation.

### **Summary of methodology**

The methodology will incorporate a mixed-method design. The research will include a document review and key informant interviews (in New Zealand) to examine the context and rationale for a country's M&E approaches and the associated results frameworks. Case studies will be selected from different contexts. The principles and approaches used, and the progress made will be examined. A thematic analysis will be completed.

The case studies in New Zealand and internationally (indicatively Papua New Guinea) will include document scans of national-level M&E reports and face-to-face interviews with national agencies including Treasury, SSC (or equivalent), Department of Planning, M&E (or equivalent), key government departments (i.e. health, education, transport, law and justice) and donors.

## **Ethical issues in relation to this research project:**

### **Recruitment and access to participants**

The recruitment of interviewees will be discussed with the chief executive or another senior manager of the organisation. It is expected that the permission to contact their employees will be granted after an internal discussion on the aim of the research. An information sheet on the research will be provided by the researcher to assist in this process.

### **Obtaining informed consent**

An informed consent process will be followed when potential interviewees are contacted. This will involve a written consent form for interviewees to sign, outlining that their participation is voluntary, that they may cease at their request, and that by signing, they agree to the digital recording of the interview. They will be sent this form and a one-page outline of the research in advance of the interview. Where it is more appropriate culturally, verbal informed and voluntary consent will be requested. Particular care will be taken during the research to safeguard the confidentiality and emotional safety of research participants.

### **Anonymity and confidentiality**

All interviews will be allocated an ID number, and this will be used during the research process. The information provided by the interviewee will be confidential and the reporting will be written in a style ensuring that no person is identifiable.

### **Potential harm to participants/researcher/university**

The questions will be asked in a professional and sensitive manner. If people appear uncomfortable during the interviews, they will be asked if they want to skip a question or end the interview. A strengths-based approach will be used in the analysis and reporting of the case studies. The style of reporting used in the case studies will ensure that organisations or participants are not directly identifiable. Care will be taken to ensure no harm will be caused by this research to participants, organisations, the researcher or the university.

### **Handling information/data**

The interviews will be digitally recorded and transcribed. A confidentiality agreement will be signed by all the transcribers and research assistant.

### **Security/privacy of information**

The information will be stored on a secure computer and any printed material will be stored in a locked cupboard.

### **Use of information**

The information will be used for the research project and associated publications only.

### **Promising access to information**

All participants will be provided with a two-page research summary once the analysis and reporting are completed. Access to the PhD thesis will also be provided to participants by an email link to the researcher's website and the Massey University library (to be confirmed).

### **Conflicts of roles**

The researcher is not currently working directly for any government or associated agencies included in this research. The researcher has worked in areas related to this research for a limited number of public sector agencies in New Zealand, donors and Managing Contractor companies in Papua New Guinea. The position of this PhD research for academic purposes will be made clear to all stakeholders and participants.

There are no conflicts of roles on behalf of the researcher.

### **Use of research assistant(s)**

A New Zealand-based research assistant will be employed to set up interviews and communicate with participants as required. This will help the researcher to confirm and conduct interviews in Papua New Guinea, where internet access and other communications can be intermittent. A support person with local language and knowledge of local culture will assist in the international case study.

### **Cultural/gender concerns**

Care will be taken while working with different cultures and genders. The researcher is experienced in working in these countries as an evaluation consultant.

## **Appendix D: Data collection tools**

- I. Key informant interview guide (initial)
- II. Updated – interview guide case study
- III. Analysis template statement of intent/annual report.
- IV. Case study – impact model results rating table.

## **I. Key informant interview guide (initial)**

Kate Averill is undertaking PhD research on the Country and Sector Results and Outcomes Frameworks

### **Aims of the research:**

The aims of the research are:

1. To research the principles, approaches and practice currently used for developing results and outcomes frameworks in different contexts.
2. To research the changes in evaluation principles, approaches and practice from the increased emphasis at country and sector levels and associated use of results and outcomes frameworks.
3. To identify the key concepts, approaches and practice for the architecture of country and sector level results and outcomes frameworks.
4. To research the impacts and significance of changes in the architecture of results and outcomes frameworks resulting from the increased emphasis on country systems and the use of frameworks for development and aid effectiveness and the implications for countries, donors and practitioners.

### **Aim of the interviews:**

The aim of the key stakeholder interviews is to build the understanding and knowledge base of the topic from different perspectives.

Details:

- The interview will be for up to 45 minutes
- Consent form to sign
- Voluntary, and can stop at any time
- Recording, transcribing, sent back for verification
- This interview guide contains questions covering the research areas.

## **Section 1: Introduction**

Please introduce yourself – role, how long at organisation

1. What is your involvement with results and outcomes frameworks, monitoring and evaluation, strategy within your organisation and in the wider environment?
2. What is your understanding of how (your) organisation's or sector's results or outcomes framework were developed?
3. What do you see as important (principles, approaches, etc.) for results and outcomes frameworks?
4. Are there any comments you would like to make on development, outcomes, strategy, monitoring, evaluation and programme architecture that you feel are important to be aware of or consider in this area?

## **Section 2: Results and outcomes frameworks – principles and approaches**

5. What is your understanding of a 'results or outcomes framework'?
6. Why are such frameworks they being developed?
7. What is their purpose? Use?
8. What level (sector, agency, programme and project) do you see as the main focus at present for strategy, planning, monitoring and evaluation?
9. What influence do you think a focus at country and sector level will have on approaches and practice to development and evaluation?
10. What principles and approaches do you see as important in the architecture and use of results and outcomes frameworks in different contexts for the future?

## **Section 3: Practitioners – development and evaluation**

11. How would you describe the current theories, principles and approaches used by development and evaluation practitioners in the architecture and use of results and outcomes frameworks?
12. In what ways do you think practitioners are changing their evaluation theory and practice when working in their own countries and in other countries?
13. What do you see as the key principles, theories and approaches for practitioners when designing and using results and outcomes frameworks?

#### **Section 4: Impacts on development and aid effectiveness, and implications for countries and donors**

14. What do you see as the role of the private sector in country development?
15. Where do you see the role for donors in country development?
16. How do you see the alignment occurring between country, sector, management, and donor programmes in country and sector development? Role of frameworks?
17. How do you see contribution and attribution from different stakeholders occurring?
18. How do you see balancing the ownership of country-focused development, country and donor accountability and capacity?
19. What do you see as the impacts for development and aid effectiveness from focusing at country and sector level and using results and outcomes frameworks?
20. What do you see as impacts of a country and sector focus and use of results and outcomes frameworks on economic development and self-sufficiency?
21. Do you have any comments or further information you would like to add?

**Thank you.**

Next steps – transcribe, send back for verification and confirmation.

You will receive a summary of the findings and papers written as research progresses.

## **II. Updated – interview guide case study**

Kate Averill is undertaking PhD research on the Country and Sector-level Results and Outcomes Frameworks

Aims of the research:

The aims of the research are:

1. To research the principles, approaches and practice currently used for developing results and outcomes frameworks in different contexts.
2. To research the changes in evaluation principles, approaches and practice from the increased emphasis at country and sector levels, and the associated use of results and outcomes frameworks.
3. To identify the key concepts, approaches and practice for the architecture of country and sector results and outcomes frameworks.
4. To research the impacts and significance of the changes in the architecture of country and sector results frameworks from the increased emphasis on country systems and the use of results and outcome frameworks for development and aid effectiveness, and the implications for countries, donors and practitioners.

### **Aim of the interviews:**

The aim of the interviews is to build the understanding and knowledge base of the topic from different perspectives.

Details:

- The interview will be for up to 45 minutes
- Consent form to sign
- Voluntary, and can stop at any time
- Recording, transcribing, sent back for verification
- This interview guide contains questions covering the research areas.



## **Introduction**

Please introduce yourself – your role, how long you have been with this organisation.

### **Development, evaluation and frameworks**

1. What concepts and principles do you consider important to be aware of for development, strategy, results/outcomes, monitoring, evaluation, management activities, programmes and projects?
2. What is your involvement with results and outcomes frameworks, monitoring and evaluation, and strategy within your organisation and in the wider environment?
3. What is your understanding of how (your) organisation's or sector's results or outcomes framework were developed?
4. What do you see as important for countries, sectors, organisations, donors, programme managers, and development and evaluation practitioners to consider in the architecture and use of results and outcomes frameworks? At the higher level? At programme level? Why?

### **Impacts and significance and implications:**

1. What changes do you see happening in development and evaluation? (If any, continue to question 6).
2. What are the impacts and the significance of these changes for:
  - a. the architecture and use of frameworks
  - b. evaluation practitioners
  - c. countries, sectors and donors
  - d. Management and governance?
  - e. Other:
3. Are there any other comments you would like to make?

### III. Matrix for analysing statements of intent/annual reports in public sector organisations

This matrix (next page) was used in each case study to analyse the architecture and use of frameworks. The criteria were selected based on the concepts emerging from the literature reviewed for this research (Appendix A) and the diagram displayed in Appendix B.

**Table 17. Analysis template for public sector agencies/sectors**

Criteria	Rating					
	Excellent	Good	Fair	Poor	Not evident	Not relevant
1. Context analysis						
2. Needs assessment						
3. Higher level framework in place						
4. Theory of change for higher framework is clear						
5. Baseline data is evident for results/outcomes						
6. Needs of different population subsets (NZ Māori, Pasifika, etc.) are being worked towards						
7. Trend analysis is based on robust data						
8. Alignment of outputs from budgets to outcomes is clear						

Criteria		Rating				
In place:	Excellent	Good	Fair	Poor	Not evident	Not relevant
9. Aggregation of data from activities and programmes						
10. Activities and programmes are aligned to higher framework						
11. Intent of activities and programmes is clear						
12. Theory of change/action for programmes is clear						
13. Programme data is evident						
14. Use of and learning from evidence are clear						
15. Agency/organisation improvement is being tracked						

## IV. Case study – impact model results rating table

Name:

Date:

Kate Averill, as part of her PhD research, is researching the impacts and significance of changes in the architecture of results and outcomes frameworks and the progress made in enhancing development and aid effectiveness, management and governance.

### **There are 3 questions to this research**

Q1. Which of the following best describes your involvement with results and frameworks, development and evaluation? *Please circle the one that best applies.*

1. Evaluation personnel
2. Management personnel
3. Donor
4. Other

Q2.From your perspective, please circle the number for each impact to rate the progress made within your organisation and the associated services/target groups.

<b>Impacts (from impact model, Figure 9)</b>	<b>Consistent progress evident</b>	<b>Some progress evident</b>	<b>No progress or change evident</b>	<b>Unsure</b>	<b>Not applicable</b>
1. Increased use of needs analyses	1	2	3	4	5
2. Increased use of results frameworks for national development strategies and sector programmes	1	2	3	4	5
3. Increased focus on country/sector systems	1	2	3	4	5
4. Increased volume of aid flow aligned to national priorities	1	2	3	4	5
5. Increased amount of aid is untied	1	2	3	4	5
6. Improved ownership of development	1	2	3	4	5
7. Improved clarity of sector development goals	1	2	3	4	5
8. Increased programme planning to meet priorities	1	2	3	4	5
9. Improved capacity in management	1	2	3	4	5
10. Improved management processes	1	2	3	4	5

<b>Impacts (from impact model, Figure 9)</b>	<b>Consistent progress evident</b>	<b>Some progress evident</b>	<b>No progress or change evident</b>	<b>Unsure</b>	<b>Not applicable</b>
11. Improved capacity in M&E	1	2	3	4	5
12. Improved M&E processes	1	2	3	4	5
13. Increased linking of priorities with expenditure framework and budgets	1	2	3	4	5
14. Improved operational development strategies	1	2	3	4	5
15. Improved results from activities and programmes	1	2	3	4	5
16. Enhanced evidence-based decision-making	1	2	3	4	5
17. Increased accountability	1	2	3	4	5
18. Increased development effectiveness	1	2	3	4	5
19. Improved aid effectiveness	1	2	3	4	5
20. Improved lives for people and environment	1	2	3	4	5

**Please comment on:**

- a) Current changes you consider are impacting on development and evaluation:
- b) Key concepts to enhance the architecture and use of results and outcomes frameworks:
- c) Implications for evaluation practitioners, management and governance:

**Q3. Other comments:**

## Appendix E: Sample of qualitative findings analysis

Contextual layer: Ideas Interventions Frameworks	
Architecture of ideas, interventions and frameworks contextual layer	
Theories of change, outcomes, systems contribution attribution	<ul style="list-style-type: none"> <li>• <b>Levels of frameworks important</b> – strategic, programme, projects/activities. Alignment is important. Strategy driven rather than policy to ensure alignment. Show links between multiple agency spend and results with contribution to longer term outcomes to achieve social and economic goals. These sector frameworks are overarching for multiple agencies and Crown entities. Sector-wide assist towards considering sector outcomes. Considering how to aggregate results and contribution to strategic level of framework. Levels of frameworks useful and contribution of providers enabled from programmes. However greater clarification for aggregating data. required.</li> <li>• <b>Need to link programme theory at top at how change is going to be achieved and how to lever that.</b> Need to prioritise and have monitoring data. Not routine activity to map interventions. By having clear strategy makes it more explicit with multiple initiatives which are linked to plan.</li> <li>• Design more for programme upfront</li> <li>• <b>Using intervention logic to link programmes projects to agency and sector frameworks.</b> To demonstrate attribution, need to get measurable impact. Important to unpack middle layers and identify what information we need to report against. Use of logic models to show contribution to higher outcomes.</li> <li>• Resourcing and capability important. Links to fiscal contributions and interventions. Started at bottom to map to higher framework. Use this for monitoring.</li> <li>• <b>Important capability at higher agency – Ministry lower levels</b> e.g. schools and kura. Need to map what people are doing and show contribution at different levels in business planning and reporting</li> </ul>
Values collaboration consultation	<ul style="list-style-type: none"> <li>• <b>Has consultation collaboration taken place internally?</b> Is it forward focused? Requires a reflective considered strategic process.</li> <li>• <b>Values are important and are often invisible.</b> Challenge with different values of stakeholders and need transparency in how value is reached. Frameworks help transparency are more explicit can be debated and contested. Identify assumptions. <i>Think frameworks should be explicit and explain how things are valued. Those things are really important for your whole framework values are often invisible</i></li> </ul>



## Contextual layer: Ideas Interventions Frameworks

	<ul style="list-style-type: none"> <li>○ It is challenging and that's why you can always debate what is the right performance measure, how do we know we're being effective and I think that's the challenge for evaluation as well because evaluations of programmes are often long term they have multiple, there's always multiple stakeholders involved, they all come with different values and so there is never a measure,</li> <li>○ There is never a measure that's important to different routes and how do you actually get some transparency in the way you approach an evaluation or performance measurement</li> <li>○ The good thing about frameworks is that they are an exercise in trying to be a little bit more transparent it's saying if you're using this framework, we would measure performance in this way but there are other frameworks which may value different things. and how does your framework, how can you be transparent in the way you present your framework.</li> </ul> <ul style="list-style-type: none"> <li>● <b>Frameworks makes things more explicit and more able to be debated and contested.</b> The good thing about having a framework and trying to use a framework is that you actually make it explicit about how you're valuing things.</li> <li>● You make gaps in your evaluation framework and performance framework visible, but I don't think frameworks should be fixed but I think they should be explicit.</li> </ul>
Change of state context processes and outcomes leadership	<ul style="list-style-type: none"> <li>● <b>Process is important, focus on changes of state rather than institution</b>, involvement of key stakeholders, show contribution. The process is important where the sector frameworks are developed in dialogue workshops, and the alignment of agency outcomes is part of this process. Once the sector framework is in reasonable agreement and approved by Boards, the agencies' can then display sector outcomes with their own agency outcomes aligned.</li> <li>● Focus on key drivers to achieve thriving sector – social and economic</li> <li>● <b>Keep to three layers for outcomes trends top, underneath key outcomes – changes of state.</b> Underneath from outputs. There needs to be more joining up using outcomes frameworks and more accountability.</li> <li>● <b>Links between results/outcomes framework and accountability – how to show contribution and context is important.</b> Need to align levels of frameworks and show intended and actual contribution and attribution from programmes.</li> <li>● There are different levels of performance reporting in agencies. There needs to be a line showing attribution for direct outcomes and contribution to shared outcomes at higher level. Then there is the higher level of wider environmental government/state level. Need quality measures and focus more on quality and affecting change.</li> <li>● Output plan is insufficient for accountability. Outcomes thinking within organisations important and need to make decisions based on outcomes. Link up frameworks and plans.</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

	<ul style="list-style-type: none"> <li>• <b>Importance of leadership from chief executive level. Required guidance</b> for linking Statement of Intent outcomes to outputs and actively articulated.</li> <li>• Outcomes for Māori and Pacific Peoples need to include.</li> <li>• Role of evaluators included in working groups for frameworks makes it more robust.</li> </ul>
Using portfolio approach	<ul style="list-style-type: none"> <li>• Top down/bottom up. Identify key benefits to lighten framework – rapid and common sense what's needed what's contributed and where do people fit in?</li> <li>• Identify key priority areas knowing what other agencies are doing and identify and segment for specific groups e.g. Pacific communities.</li> <li>• Gather info on impacts and how their lives have changes as results of programme strategy.</li> <li>• Changed structure within agency to work on theme i.e. community, education etc.</li> <li>• <b>Have formative / design phase</b> to document then link doing monitoring and evaluative outcomes.</li> <li>• Needs to be sector driven rather than institutional driven align business units <ul style="list-style-type: none"> <li>○ We also developed this new constructive in the agency and have theme teams related to each of those four areas.</li> <li>○ we need to identify key priority areas that we will focus on, taking in to account and consideration the size of the ministry and current spend and budget that we work within.</li> <li>○ Once we have our clearly defined priorities, and just year we underwent a whole another scoping exercise to do that for the ministry. And we have four key theme priority areas, one of them is education, the other is community and social enterprise. One other is youth and skills and the last is languages and culture.</li> <li>○ Within each of those four theme areas, our broad priority areas, there are various activities that are happening between the relationships you're working with the communities and the policy team</li> <li>○ Obviously, a challenge can be where a need is felt in trying to develop a logic around it Logic can be a bit hard to unpick. if a big issue blows up and you put in place, which is fair enough, the issue is what's the real issue here? So what's the best way of approaching it and I think it's important they deliver the initiatives that they need to. If it's been in response to something that's happened, but also addressing what the underlying issues might be as well. That a bit of a harder challenge</li> <li>○ Within this organisation is after a formative phase where that can be cause but if you're lucky there's a formative phase. We have evaluation plans. You'd look at how it was set up with logical model... how they work in practice and the monitoring long term outcomes in the initiative.</li> </ul> </li> </ul>

Contextual layer: Ideas Interventions Frameworks	
Role of evaluation in design	<ul style="list-style-type: none"> <li>• <b>Strategy and planning team role on frameworks and accountability.</b> Identify areas of contribution i.e. road safety linked to what is being done.</li> <li>• <b>Role of evaluators included in working groups for frameworks makes it more robust.</b> Include evaluators and have one overarching framework and can show shared and different contribution of different social sectors. Keep it simple and include people, capability and values. Need quality measures and focus more on quality and affecting change.</li> <li>• Evaluate strategic choices and adapt at different levels strategic or lower levels.</li> </ul>
Adaptive iterative	<ul style="list-style-type: none"> <li>• Need to identify things as you proceed. Will not know all indicators.</li> <li>• Planning and bottom up feedback on what information need at moment, and identify areas want to measure – iterative process.</li> </ul>
Subgroups – ethnicity, gender and regions	
Inclusion and participation	<ul style="list-style-type: none"> <li>• <b>Participation important. Inclusion of different cultures</b> Māori iwi in programmes, other cultures Pacific Islanders</li> <li>• Include in frameworks approach what's driving slightly different but same</li> <li>• Moving from grievances of past and settlements to going forward what's needed. Turning point going forward what's required.</li> <li>• Making sure Māori are part of society – social and economic. Māori succeeding as Māori. In history Māori was not taught.</li> <li>• Māori Development Agency (TPK) to influence other ministries to make sure Māori considerations is in their hands and Pasifika and Asian. For Pasifika include in frameworks priority areas.</li> <li>• Inclusion and contribution of Māori within sectors. Requires a mechanistic linkage to make it happen. <i>Some do things communicating not meaningful way not enough time take into consideration effect</i></li> <li>• <b>Allocate budget look at needs of communities.</b> Dialogue with communities what's happening what's needed</li> <li>• <b>Work with other govt departments to identify gaps where priorities are and to address gaps.</b> Important to know what other agencies are doing. Link on policy front how better to enhance outcomes. Work across sectors. Use evidence to show gaps longitudinal studies</li> <li>• Context important broad then specific. Share knowledge and complementary areas.</li> <li>• Need mandate so has to happen community, policy, operational, performance and reporting. Want to engage at outset currently ad hoc. Plan together and undertake meaningful way. This would result in more impact stronger relationships.</li> </ul>

Contextual layer: Ideas Interventions Frameworks	
Groups, indicators and targets, segmentation & reporting	<ul style="list-style-type: none"> <li>• <b>Using systems approach to target groups within.</b> Include Māori dimensions in indicators Within outcome frameworks have specific outcomes and targets Importance of Treaty and strategy of Māori Pacific with agency. Does not appear systematic. Once over lightly but not involved in really driving change</li> <li>• <b>State sector effectiveness for Māori reporting at high level but not lower.</b> Move from communication to influencing. Measuring outcomes at different levels</li> <li>• Relationships important Have longer terms partnership working together Include Māori governance and structures – still developing.</li> <li>• <b>If priorities set separate targets for outcomes and indicators specials needs cultures, gender.</b> Stretch targets to improve lives and specific interventions to targets efforts towards them. Targets groups that are not being served. Measure areas where influencing</li> <li>• <b>Use of focus agencies TPK and Pacific to focus and connect policy.</b> Role of M&amp;E– monitoring across agencies and evaluation focused within on initiatives TPK mandated to look across agencies. Focus agencies TPK and Pacific provide advice it is not clear how that works networks to other agencies' policy. Not clear of rights and roles of TPK and Pacific. Whose role is it to measure agencies as e.g. Pasifika don't have resources to measure all.</li> <li>• <b>To get more results hen work together and target priority areas</b> Then set more ambitious targets. Health looks ad different groups health issue measures. Look at disparities Māori Pacific getting better. Averages hide.</li> <li>• <b>Need specific initiatives.</b> Need to have frameworks and for youth.</li> </ul>
Regional focus	<ul style="list-style-type: none"> <li>• <b>Appears limited mention of regional priorities in NZ – more in PNG.</b> Team working with regional offices so both national and regional offices</li> </ul>
Use of Frameworks	
Sector and system driven and institutional performance	<ul style="list-style-type: none"> <li>• <b>Use layers in frameworks to assist with measurement Indicators provide data at different levels and cohesiveness for sector.</b> Use the key data that is available from concrete measurement perspective. There is a distinction between institution driven as opposed to being sector driven which provides wider results over multiple institutions contributing.</li> <li>• <b>People and priorities important.</b> Get clear on terminology. Get clear what mean terminology means guiding people on priorities</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

	<ul style="list-style-type: none"> <li>• <b>How models are used is important.</b> What they mean requires communications and knowledge. Require clarity on terminology clarify what mean by models and guiding people to interpret them.</li> <li>• <b>Sector programmer frameworks are across agencies.</b> Require clarity about what intending to achieve for outcomes and for population groups i.e. Māori. Require links between activities, outputs to outcomes and links to higher outcomes/impacts which are longer terms trends. Need to see how will make contribution to get changes demonstrate and make explicit.</li> <li>• Sector-based approach with whānau ora really pulled sectors together used to be at strategic level and then line of sight but not following up. Leadership needs to have a view and ensure incorporation need to generate work and outputs to contribute to subsidiary and up to strategic level linking in with strategic goals and asset base needs to become more obvious what doing what contributing and what results</li> <li>• Components of systems. May have variations in services some performing well but system overall not performing integration important right mix.</li> <li>• <b>Hard for people to understand where they fit in system complex and involved</b> – system view – performance of area people relate to but thinking about big picture and finance important. Funding flows and cost and people need to align. Systems approach required some parts may work but need to look at overall.</li> <li>• <b>System performance and component performance</b> Need people to be aware of bigger world and then see their part in it. Need to know where money is going and how that's getting down to the people.</li> </ul>
Values, interpersonal, dialogue key priorities contextual inclusion leadership	<ul style="list-style-type: none"> <li>• <b>Top down bottom up approach, dialogue and workshop start with draft framework</b> – key results outputs outcomes impacts Basically what we did was we used the straw man approach, so we basically built what we thought was what the sector framework might look like and put that in front of them and then worked that up and teased that around. And that was purely a tactical process because there's no point in starting with a clean sheet of paper on these things</li> <li>• <b>Emergent outcomes different perspectives view. Governance important</b> between on who can make the call. Environment body, iwi groups. Different views of units of outcomes of vested groups and owners. Different perspectives and outcomes for stakeholders inclusive</li> <li>• Formalising processes joint commitments to review assess making adjustments Country sector adjust for implementation managers involved taking information and using in decision and feeding it back for accountability learning and management processes. At country level clarity around results frameworks and accountability. At activity level less prescriptive.</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

- **Living documents and make adjustments, two parts collating information and dialogue with development partners.** Second is feeding it back to institutions for accountability learning and management processes. Adaptive processes and dialogue inclusive processes
- Dialogue and workshop Leadership important analyse reflect. Need business unit want to work together Problem-solving management need leadership important DCEs. Workshopped, helped each other good working knowledge across business units, problem solving collaboration. Leadership from senior executive important
- Dialogue contextual adaptive. Good communication contextual framing key priorities adapt *not all change is good so monitor changes reflective practice*
- After annual report bringing everyone back together again – right people in room – need to have done prep work good prep good communication ahead so can see areas to improve or not change if change where is measure where is data asked them top priorities for change not all change is good learning what happened why?
- **Use governance groups overarching and across agency external inputs assist getting out of institutional thinking contextual.** High level governance group involvement for around outcomes frameworks – particularly cross agency. Role of who leads development external led.
- **Use frameworks to get alignment priorities for funding people.** Suggest we need to structure our documents around frameworks so it makes sense have dialogues and then articulating in frameworks. Using frameworks to get feedback how work relates to different framework and then can write about it so it makes sense. The outcomes in the framework show what you are wanting to achieve. Gives guide to ministers and creates dialogue and sometimes tension. The SOI provides clarity from conversations and out what are government and ministerial priorities. Show where the funding is going to go and the frameworks assist dialogue and clarity.
- Underpinned by values which are important as informs what outcomes are important i.e. end of life disability, value for money quality different stages of life.
- How is it meeting different population segments, ethnicity using more than averages, distribution geographic, need to segment and analyse with other information income deprivation.
- **With outcomes and frameworks value x and what do we value how are we concerned at distribution** different stages of life effectiveness and value for money whose using services not assuming same look at geographical variation ethnicity deprivation. It's the accountability of delivery and how it works to create a system which is the outcomes for Ministry.

Contextual layer: Ideas Interventions Frameworks	
	<ul style="list-style-type: none"> <li>Links policy strategy business measuring sectors performance – how well they are going against plan, health of population and programmes contribution, contracted services are they right services, people quantity efficiently and effectiveness – the right mix.</li> <li><b>Dialogue and prioritise</b> Look at what prioritise over others different from audit level when accounting for all things need outcomes to become more stable as NZ focused getting what you are working to achieving and people on board is important.</li> </ul>
Strategic alignment, priorities and framework layers theory of change, programme theory	<ul style="list-style-type: none"> <li><b>Align results at different levels so can aggregate results line of sight and get some indicators.</b> Identify key results at different levels and what is important. Can map down to each level of service. Want high level trend indicators, set targets at lower levels. Strategic align, identify gaps, caveats changes of state <i>Seen high level outcome identified. Need to be real</i></li> <li><b>Supply and demand. Understand structure more and contribution policies that contribute align well</b> <i>what are the intermediate outcomes.</i> Under contribution strategic alignment. Use frameworks to identify important priorities different stakeholders level community provider institution sector. Trend analysis and look at contribution.</li> <li>Purpose of using frameworks is more than communications conceptually what is important for whānau and content tension between taskforce Cabinet paper and provider and sector context needs to articulate both so look at contribution of current activities and situation. Balance frameworks reality and show situational of taskforce and Cabinet. <i>What are you trying to measure effectiveness of sector system? Question what is the purpose of an outcomes framework? – hang measurement off different levels or communication device at lower level – wellbeing at whānau level and provide government levels – effectiveness of programme.</i></li> <li>Conceptually at whānau and provider different measures constrained by taskforce saying what outcomes framework can it conceptually stand up to be both? Limitation only driven by current data collected or get framework and then look at data and gaps. Can end up with stock standard indicators and the gaps become forgotten.</li> <li>Can starts with framework around social at high level and include relevant outcomes – then identify indicators which become proxy and can be ok. But caveats and gaps can get forgotten – require clarity on purpose of framework what's included some indicators and some more meaningful measures on well-being. If just use standard indicator don't get changes of state provide for some ongoing and more meaningful measurement if you like.</li> <li><b>Strategic alignment and reporting. Need direct mapping and show what is and what has done – system and contribution.</b> Annual reports can be just a list of initiatives but no mapping of direct contribution. Needs to be more joining</li> </ul>



## Contextual layer: Ideas Interventions Frameworks

up from intervention to outcomes frameworks for accountability. Strategic alignment, reporting and accountability. Thinking more about outcomes correlation and agency accountability important.

- With outcomes need to be thinking of correlation and contribution of interventions and accountability for those outcomes changes positive and negative so can look to address adapt with changing contexts.
- **Use frameworks to discuss conform priorities align results.** Keep it simple big things in frameworks sectors. With the framework – social sector partners need to be able to see where they fit. This was a focus upfront. To get simple framework took months but once there it makes measuring and reporting outcomes that much easier.
- Identify line of sight contribution what was needed contribution Tension with telling contribution story and knowing what people have done. With managing for outcomes coming in, focus was on measuring outcomes, management for outcomes lost focus and measuring outcomes was sufficient without telling contribution analysis story which is important to demonstrate success.
- **Strategic alignment management processes and results** Management processes important coherent story link between quality of what do and what achieve and quality of outcomes. Moving from outputs to outcomes took time. *With management piece being able to tell coherent story around how organisation direct contributes to quality and what other people do influences the system. Link between quality of what we do and quality of outcomes still needs attention.*
- **Programme theory important** – intermediate outcomes what contributing to Have logic model or model with shorter and longer-term outcomes there is demand from policy to measure longer term outcomes in the higher-level strategic space which requires clarity. If programme theory clear then evidence is available if intermediate outcomes being achieved likely to achieve strategic longer terms and can retest in summative way later on – link between logic, monitoring and evaluation
- Programme theory is required to explain different levels strands government agency provider, social service providers, whānau strands – different outcomes sought at different level. Unpacking what different levers are at different strands, policy, big diagram would assist otherwise confusing line of sight form agency investment contributing towards higher levels to across agencies inform programme theory and monitoring going forward. Programme theory at different levels – community provider strategic impact policy link investment to inputs to outputs policies.
- **Focus on targets.** In NZ there is a lot of focus on outcome frameworks compared to Scotland that focuses on targets. In NZ there is a logic to what outcomes we are seeking to achieve which are depicted in diagram measures are hidden starting at bottom to get to top. In Scotland more open-ended about what you do.
- **Link between govt agencies important PNG portfolio approach identify goals and allow budget.** Many standalone evaluations done but not linked evaluative approaches needs to clarify intents of initiative and what trying to achieve in different communities – iterative development evolves from community inputs



Contextual layer: Ideas Interventions Frameworks	
	<ul style="list-style-type: none"> <li>• Three layers community/ provider Ministry operations then Ministry strategic results – goals objectives, align results framework, make sense logic between national and line agencies</li> <li>• Portfolio approach with PNG means more bottom up initiatives. Top down bottom up within sectors. Using frameworks to assist to do that combining process and outcomes</li> <li>• <b>Need to get clear about object of change link with social outcomes and whole system</b> Trying to work towards common outcomes from community provider to implementation at government level need to look at links between what delivering to what achieving and how you bring in innovation.</li> </ul>
Data needs priorities and dimensions	<ul style="list-style-type: none"> <li>• <b>Identify what data is needed, time series gaps at different measurement levels</b> – outputs, outcomes, impacts, level of accountability for services and contribution to sector and variables. Identify key data from measurement perspective. What is imp from strategic perspective and lower down Identify data needs</li> <li>• Identifying data from frameworks was first and second phases. Then can model sector with costs and changes to outputs and links top outcomes and impacts and other impacts – model cause–effect relationships.</li> <li>• <b>Identify data needs</b> Look at where gaps are for data, processes important look at cause and effect i.e. justice and three strikes law – impact on prison population</li> <li>• <b>Data dimensions needs capturing</b> ie growth and development in education not fixed is there shared understanding.</li> <li>• Identifying results outcomes along way important but not full picture. Clarity on outcomes required and interpretations, not fixed or equally shared understanding <i>can get caught between indicator and outcome as indicators don't capture all. Total judgement outcomes can be misleading e.g. qualifications and human development human growth. Qualifications can be proxy for growth, but other things went on and changed. Changes of state.</i></li> <li>• <b>Measuring focus on outcomes and transparency on what's actually being achieved from taxpayer funds and learned is important.</b> Measuring evaluating to understand more. Clarifying the business model, measuring to know how its progressing/being achieved.</li> <li>• A transparent evaluative culture and what is being achieved is the important aspect not focusing on management process – outputs. Creating impacts outcomes changing state of things that is the business.</li> <li>• <b>Public servants need to know the business they are in converting outputs to outcomes and impacts</b> the macro service model for govt requires clarification as current one is not working requires transparency outcome focused need to have achieving outcome and evaluative dialogue and evidence in same space focus on management rather than achieving outcomes need to focus on eh business you are in creating impacts Evaluative management</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

	<ul style="list-style-type: none"> <li>• <b>Have some measures and then benchmark with other judications</b> – statements of intent says why doing it simple story and show outcomes. Can have measures for international comparisons.</li> <li>• <b>The body of statement of intent tells story of what ministry does.</b> Technical legislative information is in appendix Keep as simple as possible with targets measures in appendix.</li> <li>• Use frameworks targets for groups measures</li> <li>• With Pasifika needs more emphasis stronger relationship require pushing more to achieve outcomes inclusive in development target setting. <i>Use logical frameworks to guide activities and can align what is going to make a difference to the mix. Use to influence discussions and reviews and targets – are they sufficient to make a difference. This way of working is embedded into ministry frameworks e.g. Ministry of Education.</i></li> <li>• <b>Measuring different levels.</b> Use indicators, higher SOI ones at different levels and then lower down from different perspectives and realities – look at how measuring outcomes – different types of indicators</li> <li>• Different perspectives in data sources. Lower level frameworks that link to SOIs important as that is where real outcomes are that you are trying to achieve. Need performance internally and perspectives of beneficiaries then out to regional levels to compare region to region.</li> </ul>
Value for money portfolio approaches	<ul style="list-style-type: none"> <li>• Value for money looks at outputs costs outcomes impacts and judgement call. Different in different countries with different policies social drivers etc when trying to benchmark</li> <li>• Value for money what's needed – why ground up rather than pots of money distributed. High level outcomes come decided by govt of day. Pots of money to achieve outcomes.</li> <li>• Two paradigms operating to meets the needs of government and people what is important – accountability but must meet needs people on ground so they would like some of that and apply – need to ensure asking correct data and that they have resources to deliver on what is needed</li> <li>• <b>Bottom up top down meet in middle.</b> In NZ have scenario with big outcomes that machinery of government wants to happen and programme and policies. Appears to be mismatch is it working? – two paradigms – how do you know those pots of money are making a difference? Have people on ground saying were doing this and we will have some money for that. It needs to meet in the middle. How are they actually contributing to those big outcomes on the ground – community organisations and providers. Have to be careful that they community providers provide key data - not take up huge amounts of their time. Need to make sure community organisations etc are accountable and realise they are doing more than what is reported.</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

Adaptive design, iterative frameworks, funding, measurement, reporting and learning

- **Need to ensure inclusion of other sector results for structural balance.** Identify indicators may change little over time as emerge. Have series of “scorecards” of how impacts from services and the feed through to higher levels to ensure measurement side is right.
- **Activities are funded through outputs. Redesign over time to get the benefits.** Drivers are the changes you want to make and politics. Starting to use sector base budgeting. Have to look at services and contribution may need to adapt model and series over time to get desired results
- Back track as want services you need to achieve – Adaptive design. Adaptive frameworks Combine social cultural economic environmental use by managers. Under frameworks economic sector has reasonable indicators, social evolving and environmental lagging.
- Role of managers important in terms of policy priorities and decision-making. Active use – 6 months 12 months Active use, staff and wider dialogue in discussion.
- What needs to change Use with partners answering key questions do together. How are we tracking? Are we making progress? What does need to change or tweak or do things?
- **Using for learning, accountability, reporting and holistically.** Accountability and learning needs to be focused on outcomes holistically. *Real time embed result you are working towards and value of them for people and environment not just accountability but value system performance and agency performance*
- **Purpose and role of SOI documents – allocate funds, identify results.** Get judged on what doc looks like. What has been achieved is not focus – assumption is that if these documents with govt depts. Are focused on outcome – are they are achieving outcomes? Is that really the case? Used to think would know in ten years at end outcomes, thinking have info earlier thinking, more holistically current approach of putting system and ministry performance together is not very elegant.
- **Adaptive Use in iterative nature, contextual links and assumptions.** Iterative development over first year. Useful and refining to get increased clarity. Strategic alignment from bottom up available resources to grow capability and to embed cultural social dimensions
- **Logic models constructed not yet being used in an iterative manner and adapted to programmes with change and as evolve.** As get to evaluation 2–3 years later not helpful need to record assumptions, links, levers, theory what was intended to lead to another.
- **Improve people capability and wellbeing managing initiatives increase capability.** Kaitoko Whānau – family advocates requires capability at community, whānau and provider and for sport good examples of embedding integrated

Contextual layer: Ideas Interventions Frameworks	
	practice supported by programme manager iterative refining over first year Wanted to improve capacity internally – involve programme people in discussions and evaluation. Support programme initiative. Involve in strategic redirection of funds and how will be used involved in decisions on how funds are used with key stakeholders including those who have had funds before using outcome framework ground up and aligning to higher framework. It makes programme manager's job easier as clarifies programme. Not top down involved managers as well going back and thinking with stakeholders.
Capability leadership managers	<ul style="list-style-type: none"> <li>• <b>Results based approaches useful.</b> Need to develop skills and strengths in business unit on how to evaluate and measure outcomes different business units do things in different ways – need to standardise approaches and how measure as social sector tend to develop own logics in own unit and people don't know about it developing skills and stents.</li> </ul>
Measurement	
Data collection systems	<ul style="list-style-type: none"> <li>• <b>Get data systems designed and links with bigger data. Different levels of measurement and reporting within system.</b> Key results outcomes agreed on Linked with assumptions, funding mechanisms, policy, outcomes Alignment and use of system approach links and integration to business processes important</li> <li>• Needs collaboration and have accountability agency, provider whānau levels.</li> <li>• Use microdata systems and collect qual data and numbers and link. Need internal systems for collecting data. <i>Question over what is sufficient?</i></li> <li>• <b>Needs to have more interest in use.</b> Access to data use protocols important</li> </ul>
Demand and need	<ul style="list-style-type: none"> <li>• Know need and demand and then move into measurement</li> <li>• Need to know how will analyse and have process. Need to define success story</li> <li>• Need to look at variability with different groups.</li> </ul>
Frameworks, outcomes and impacts	<ul style="list-style-type: none"> <li>• Outcomes theory of change links changes in state. Need to establish links between policy and outcomes and impacts. Align goals objectives outcome targets</li> <li>• Need to use framework to guide reporting data collection and aggregation from multiple providers</li> </ul>

Contextual layer: Ideas Interventions Frameworks	
	<ul style="list-style-type: none"> <li>• <b>Measuring much clear easier with framework. Use outcome frameworks and logic models so align.</b> Measure but need to influence i.e. Pacific Islands – how much influence. <i>How much influence are interventions having?</i> Consider baselines and changes over time When do framework think about types of indicators need, process outcome impact</li> <li>• Role of programme manager to align and include evaluator with skills</li> <li>• Social outcomes frameworks link with SOI Need to look at indicators in SOI for relevancy and <b>measuring different perspectives</b>. High level alignment ok Making judgements how are progressing what are contributing to.</li> </ul>
Data sources, security and non-financials	<ul style="list-style-type: none"> <li>• <b>Alignment and use of system approach links and integration to business processes.</b> Auditing for non-financial information is not that different to financial – look at data source integrity and controls over data systems in place.</li> <li>• Measurement is for different audiences – populations, agency provider whānau.</li> </ul>
Indicators and targets	<ul style="list-style-type: none"> <li>• Use frameworks to guide measurement. Measure at different levels funding processes outputs outcomes and impacts align with funds. Identify meaningful outcomes. Measure at different levels. Use big data and specific data with line of sight so can assess contribution.</li> <li>• Measures roll up into strategic direction and where agency is going. Also align indicators cross agencies.</li> <li>• <b>Establish targets for different populations groups.</b> Include trend analysis not always against target. Establish a baseline</li> <li>• Adapt as you go – change inputs to achieve outcomes and impacts.</li> <li>• Track over time so get sense of improvements will be incremental progress – move beyond baseline stage. Context links causes of numbers.</li> <li>• Some indicators re performance and some need to be outcome Need different indicators to get what is being treated by interventions. Be clear on what you can and can't attribute</li> <li>• <b>With new initiatives establish measures and baselines.</b> Contribution to health measures e.g. important can show making a difference</li> <li>• <b>Need right information being measured.</b> Have hierarchy of indicators but not necessarily the indicators that we would need to measure our progress against outcomes. Robust and credible and purposeful. Analysis important making sense of data.</li> </ul>

Contextual layer: Ideas Interventions Frameworks	
Qualitative and quantitative data collection	<ul style="list-style-type: none"> <li>• <b>Confirm measurement priorities and key information needs.</b> Look at sample size all or sample for measurement. Goes into policy space evidence Could it have similar impact</li> <li>• <b>Regional data collection local and regional important to know what is happening</b> Need to align data collection programmes regional and role of national and provincial governments.</li> <li>• <b>Layers of result Linking NGOS and provincial data</b> Provide more structure for data collection programme regional need to monitor outputs processes and outcomes so know what is happening</li> <li>• Numbers and qualitative perceptions matter monitoring needs to include both. Tendency to focus on implementation</li> <li>• <b>Clients do measurement for themselves</b> Access to data and collecting important NGOS churches Programme data collection important</li> <li>• <b>Need to understand context assumptions replicable.</b> Time follow up capture impacts down line. Need to link and follow up. Taking a longer-term view</li> </ul>
Learning	<ul style="list-style-type: none"> <li>• Use evidence needs to be fact based and data taken from multiple sources. <i>Look at what you are trying to measure.</i></li> <li>• Information systems and links.</li> <li>• Need partnerships to get outcome data</li> </ul>
Role of M&E	<ul style="list-style-type: none"> <li>• Role of managers. Different capability and background with measuring outcomes Track and measure programme is programme manager's responsibility. Try and minimise burden on system and government - not replicate.</li> <li>• First step is pulling together framework so know what is important. Design around evaluation models – important pragmatic and can inform policy.</li> <li>• Value of different people being involved. Need to know what is important for monitoring and know more about</li> <li>• Clarify what do in monitoring and then in impacts study.</li> <li>• Evaluations interface with indicators. Value of monitoring and linking to measure other areas. Changing of role of evaluations used to be external now Evaluation studies more for organisational learning.</li> </ul>
Reporting and accountability	<ul style="list-style-type: none"> <li>• <b>Non-financial information is improving moving to outcome important to establish outcomes measures.</b> Issue grades on non-financials. Auditing for non-financial information is not that different to financial – look at data source integrity and controls over data systems in place.</li> <li>• Needs collaboration and have accountability agency, provider whānau levels.</li> </ul>

Contextual layer: Ideas Interventions Frameworks	
	<ul style="list-style-type: none"> <li>Monitoring economic indicators does not show attribution. <b>Need to come up with some attributable indicators that show and links actions to results.</b> Need to show attribution as accountable to taxpayer. Indicators can and do change with different governments Measure and report at right time to right people.</li> <li><b>Important for get national and regional data and reporting.</b> Then down to site level – frequency of reporting needs to be considered. <b>Getting info more frequently</b> – weekly, monthly daily. Monthly can analyse more. Need to get commentary around. Weekly can have fluctuations s so monthly good.</li> <li>Link reporting from different business units. Reporting from partners important.</li> </ul>
Evaluation role and nature	
Role and function	<ul style="list-style-type: none"> <li><b>Evaluation is not regarded at strategic level, perceived at being at output level</b> <i>evaluation is not at the table.</i> Lack of understanding of role of evaluation – evaluation still sits in policy area. No consensus over role and nature of evaluation. <i>Not have clear thinking about evaluation some monitoring but need to think what we do and why</i></li> <li>Some agencies having more robust frameworks. Lack of confidence in frameworks</li> <li><b>Role of evaluation important. Aligning of monitoring, evaluation and intervention business processes.</b> Link back into business unit for monitoring and measuring impact over time. Organisation not very mature about it – need to integrate as way of working</li> <li><b>Measurement important.</b> Aligning, evaluating and proving</li> <li><b>Having a designated role for evaluation roles to focus on evaluation. Evaluation need to be mode of operation.</b> Model of operation required for evaluation. Can have internal plus some external in evaluation community of practice – common approach – don't have at present and dependent on some individuals</li> <li><b>Not have systematic methodological approach from strategy to implementers to providers so all different.</b> Some systematic approach of implementers which makes big difference but no systematic approach inside agency. No oversight</li> <li>Context specific and value for money. Some process info important so know what and why it is happening</li> <li>Integrate research evaluation and performance moving to more professional basis on providing evidence base and adapt according to size. Inform policy and provide evidence. Evaluation needs to be theory based and link to policy cycle.</li> </ul>
Nature and amount monitoring and evaluation	<ul style="list-style-type: none"> <li><b>Need evaluation involved with design.</b> Very little evaluation done small monitoring team and thinking of setting evaluation design when setting policy objectives. Internal external debate. Need to have demand for evaluation. Need to</li> </ul>

## Contextual layer: Ideas Interventions Frameworks

	<p>know what is going on currently. Information priorities. Big dollars spend and need to know there is value. Little done at present.</p> <ul style="list-style-type: none"> <li>• <b>Co-design linking management evaluation – evaluative management.</b> Linking across organisation sharing information – learning organisation. <i>Incorporate co-design with policy teams create a working environment where incorporated from beginning driven by say policy team – in their experimental approaches.</i></li> <li>• Limited capability and capacity. Culture change important to build evaluative capacity. Support managers to find out and meet info needs.</li> <li>• <b>Systems approach and role of monitoring. Use results frameworks</b> linking planning monitoring, evaluation and reporting. Use outcome frameworks, looking at governance and Audit New Zealand Treasury.</li> <li>• Embed evaluation. More than just basic monitoring. Two levels activity level – providers tracking during implementation and strategic level using results indicators.</li> <li>• <b>Flexible approach with programme managers</b> – make adjustments based on evidence. Understanding effectiveness of interventions.</li> <li>• Need glossary so all on same page and use common terms</li> <li>• <b>Collaborate right people in room discuss together with some thinking in business units before discussion.</b> Have evaluation hub, support priorities, staff placed in teams. <i>Strengths-based we are all working together – evaluation will not make boat sink.</i> Build up trust so value seen and included.</li> </ul>
Structure and view	<ul style="list-style-type: none"> <li>• Do evaluation as efficacy of things direction link with strategy and plan and have sensible measure and assess whether doing right things. Evaluators do not have reputation of connecting strategy with evaluation – evaluation focused at output level – needs to move up. Evaluators and evaluation at outputs level and do not reputation of being people who can connect strategy with evaluation. Incorporate output class as evaluation priorities. Need to understand what means not just for research but performance – significance. Current many evaluations if do at all not linked to bigger framework. Needs to be coordinated and prioritisation.</li> <li>• <b>Need to have integrated approach strategy policy frameworks work together important t to develop and have shared understanding.</b> Set up in beginning rely on theory of change set up at beginning and then monitor and measure evaluative questions. <i>How is it going against theory of change logic of initiatives?</i> Need outcome framework – systematic behaviours and perspectives important to be included.</li> <li>• <b>Cost of evaluation one reason for lightly done. Relying more on external bodies to come in rather than internally undertaken.</b> Clarity is required between policy reviews and evaluation. Business unit responsibility. Cross sector team</li> </ul>



Contextual layer: Ideas Interventions Frameworks	
	<p>MSD initiatives using frameworks commissioning etc. <i>Consider all initiatives should have some evaluative activity try and do so if new innovative.</i></p> <ul style="list-style-type: none"> <li>• <i>Evaluation that is most valuable is from learning perspective rather than compliance driven.</i> Looking at role of action research and programme policy evaluation looking at agency effort</li> <li>• Focus evaluation capacity on understanding development effectiveness. Support experimental interventions. Link with social science. Role of evaluation quite young in NZ from business perspective. By evaluating under about value of interventions. Centre of evaluation for expertise</li> <li>• Tension between understanding from evidence scientific perspective, means need to change business practice which can cause tension of performance management or learning tools – managers need to adapt.</li> <li>• <b>Be collaborative involve providers Māori evaluators understand value of collaboration.</b> Be explicit add value Having evaluative discussions conversations. Use evaluative discussion within agencies. Role of evaluators facilitator Don't overcomplicate don't make it too big. Big ticket items.</li> <li>• <b>Has to have willingness to undertake in structured way.</b> Use tools create evaluation hub use different strategies with groups and levels. <i>Mismatch between inclusion community and providers with government can learn from each other need to bridge paradigms or change paradigm so inclusive.</i></li> <li>• <b>Understand context.</b> Be responsive to change context. Need to approach creatively and dynamically e.g. approach diabetic in more embedded way use evidence and form other jurisdictions.</li> </ul>
Performance reporting	<ul style="list-style-type: none"> <li>• <b>Cross cutting reporting at three levels – provider strategy and service delivery.</b> Performance reporting research monitoring evaluation including value for money type exercises.</li> <li>• Look at provider reporting and interface with action research evaluation and overall higher government level – <i>is it working?</i></li> <li>• Look at how to measure and report systematically with clarity. <i>Identify priorities for measurement and change if required.</i></li> </ul>
Individual capacity: Evaluation practitioners' capacity and capability	
Personal attributes	<ul style="list-style-type: none"> <li>• Approachable and people feel comfortable sharing, build up trust. <i>Be able to work together</i></li> </ul>

Contextual layer: Ideas Interventions Frameworks	
Training and development	<ul style="list-style-type: none"> <li>Conflict over building capacity and doing inside agencies. Build up capacity and capability across organisations. Link to policy.</li> </ul>
Capacity and capability: skill and understanding of key business and evaluation concepts	<ul style="list-style-type: none"> <li>There is a lack of business model knowledge on business of outcomes not just on process. Limited understanding of outcomes. Need to understand what business in converting taxpayer funds to impacts need to understand this business model.</li> <li><b>Very low base of technical skill for dealing with outcome and performance information.</b> Evaluator can work from top down to details which is needed to construct framework. Operations persons are difficult to engage as need to have shared understanding. Need to focus on valuing <i>the ability to squarely and rigorously assess how an organisation's tracking over time</i></li> <li><b>Understanding difference between research and evaluation not same.</b> Evaluation part of policy system and organisational understanding of influence and impacts. <i>Need skills as base inside agency share and have community of practice Don't have people in house who understand and who can evaluate effectiveness.</i></li> <li>Need context interpretation and analysis and qualifications. Need quantitative and qualitative so understand context</li> </ul>
Individual capacity: Other agency personnel capability and capacity	
Devolution leadership and systematic embedded approach	<ul style="list-style-type: none"> <li><b>Need stability and good leadership to embed practice.</b> Role of leaders in agency to say go this way and link operations to outcomes.</li> <li><b>Need to get clear around monitoring roles, and interface with evaluation</b> Identify what information do need from these systems This needs to be communicated more clearly for people on the ground. <i>Keep it simple, be flexible.</i></li> <li>Build skills up of managers, programme staff, leadership with training. Agree on definitions.</li> </ul>
Culture and system within organisation and capability	<ul style="list-style-type: none"> <li><b>Culture, system, capability and sharing to recognise value not just processes.</b> Need to have system so people can learn. Issue of capability some have but not space to do it. <i>Takes time and good practice to embed across organisation.</i></li> </ul>

## Contextual layer: Ideas Interventions Frameworks

- Need staff, managers at different level offices to have higher capability. Slim resourcing compared to other teams within agencies. Require accountabilities to be clarified who's managing, who is accountable
- Need to understand business model and outcomes. M&E at different levels, using indicators and aggregate, harmonise with partners, feedback into strategic level. Increase planning and co-design in sectors using outcome frameworks to assist linking shared and aligned outcomes increase consistency and coverage. Need to understand system frameworks within link components and levels people roles and capability
  - **Some fudging of results outputs saying are achieving varied quality in outcome space.** Use of results frameworks positive as can understand what and why doing. Need guidelines and support to get usage across organisation. *Need to ask within agency what do we need to do to whether its achieving outcomes.* Unified way of working sharing has to make the links of how what people does contributes to strategy and outcomes at higher level.
  - **Evaluation capacity inside organisation to support managers under resourced planned.** There is some existing evaluative capability inside agencies that can be utilised. People come from different backgrounds and also variations across agencies. Background of internal personnel generalists in agencies – need more technical expertise to support managers and internally at management level. Need more technical skills such as using excel etc performance information. Capacity and capability important from ethnic diversity need to support e.g. Pacific. Extend analyst capacity and use of qualitative to give meaning to explain impacts. Need to understand data security and meaning of data and what is showing so capability in this important If people move and lose data system falls over
  - **Tension between doing evaluation and supporting internal teams.** Role of internal teams for evaluation and research is to socialise findings within agencies – need commitment from managers/leadership team. *Need personal commitment of agency personnel to take this evidence base and evaluative management seriously. Need tenacity and willingness to get to grips with design data analysis and use.*

## Appendix F: Qualitative research findings thematic summary

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
<b>Wider Infrastructure contextual layer</b>				
Relationality	Development principles and approaches, inclusion partnership and participation	Governance	Development principles and approaches	Improved lives for people and the environment
Relationality, Contextual sensitivity, Adaptation	Government and management accountability and democracy and inclusion, accountability frameworks and people, performance, constitutional arrangements managerialism, explicit theories of change, prioritising outcomes and investment, politics, donor alignment, systematic planning and prioritisation approaches and standardisation, systems functions, flexibility, regionalisation, embedding evaluative approaches, sector and country portfolio planning and investing, aligning financial systems including NGOs private sector		Government, management, accountability, and democracy	Improved lives for people and the environment
Contextual sensitivity	Service delivery, community inclusion and evaluative feedback, iterative adaptation and results		Service delivery, community inclusion and evaluation	Improved lives for people and the environment
Relationality	Development results – focus on outcomes, statistics and data, context, wellbeing	Development – approaches, power & democracy, culture, national capacity	Development results	Improved lives for people and the environment
Contextual sensitivity	Development theory and approach, government role and community inclusion, Paris Declaration principles, community and sector inclusion, power and control		Development principles and approaches	Improved development effectiveness
Adaptation	Role of Central Agencies – leadership development, evaluation and management, state sector model, explicit function for evaluation, audit role, central agency role with performance, accountability	Central agencies	Role of Central Agencies- development, evaluation and management	Improved development effectiveness

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Adaptation	Agency views of central agencies approaches and processes, use of results and outcomes frameworks, accountability, central agency role, donors aligning to country priorities		Agency views of central agencies approaches and processes	Improved development effectiveness
Relationality, contextual sensitivity, adaptation	Accountability and alignment, values and value of outcomes, results and outcomes frameworks, inclusion, integration, performance data, resourcing, accountability, devolution, non-financial information, sector approaches, regionalisation, development planning and alignment of investment, incrementally, evaluative discussions, community involvement	Accountability – performance management, audit, evaluation, financial, country systems	Government role and government to government cooperation	Improved development effectiveness
Adaptation	Performance management systems, state sector model, managing for outcomes theory, mandate for evaluation, organisational culture for sector and institutional performance and learning, coproduction		Country development, results and role of monitoring & evaluation	Improved development effectiveness
Relationality	Audit guidance and standards for non-financial information, centralisation and devolution, mutual accountability, agency maturity, provincial mandate for coordination, administration, service delivery and M&E		Development theory and approach, government role and inclusion	Improved development effectiveness
Adaptation	Results frameworks and evaluation, integration, systematic and flexible approaches, M&E focus, learning organisation, alignment of country systems and inputs, stakeholder engagement		Development results	Improved development effectiveness
Adaptation	Financial funding for outputs and outcomes, investment, prioritisation and allocation, portfolio budgeting, leadership, manager capability, country, sector and thematic alignment		Development principles and approaches	Improved development effectiveness
Contextual sensitivity	Government role and government to government cooperation, applicability to other country systems, context, constitutions, relationships, aligning to country systems	Donors – govt to govt cooperation, role of monitoring & evaluation	Government, management, accountability and democracy	Improved aid effectiveness
Adaptation	Country development, results and role of monitoring & evaluation, country-level requirements for results and accountability, activity level requirements with alignment, investment approach, evaluation role and approaches		Service delivery, community inclusion and evaluation	Improved aid effectiveness

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
<b>Ideas/Interventions/ Frameworks contextual layer</b>				
Contextual sensitivity	Theory based evaluation, programme theory, theories of change, outcomes, contribution, attribution	Architecture of ideas, interventions and frameworks	Theories of change, outcomes, contribution, attribution	Results framework
Relationality	Values – inclusion, participatory, collaboration and consultation		Values, collaboration and consultation	Emergent
Contextual sensitivity	Devolution, leadership, programme theory – change of state, theory of change, context, processes, and outcomes		Change of state, context, processes, outcomes and leadership	Results framework, and link between results, leadership, management capacity
Adaptation	Strategic planning and emergent strategic intent, theory-based evaluation, portfolio based funding		Using portfolio funding approach	Country systems
Adaptation	Adaptive and iterative framework design		Adaptive iterative framework design	Emergent
Relationality	Inclusion, participatory, segmentation		Inclusion and participation	Ownership
Contextual sensitivity	Theory based evaluation – subgroups, indicators and targets, segmentation and reporting	Subgroups – ethnicity, gender and regions	Groups, indicators and targets, segmentation and reporting	Results framework
Contextual sensitivity	Regionalisation, decentralisation		Regional focus	Emergent
Contextual sensitivity	Context, country and sector systems, adaptive and emergent strategy, sector and system driven and institutional performance	Use of frameworks	Sector and system driven and institutional performance	Sector development goals

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Contextual sensitivity, Relationality	Values, prioritisation, inclusion, interpersonal, dialogue key priorities, contextual inclusive leadership		Values, interpersonal, dialogue key priorities contextual inclusion leadership	Emergent
Adaptation	Theory based evaluation, strategic planning, strategic alignment, priorities and frameworks – layers, theory of change, programme theory		Strategic alignment, priorities and framework layers, theory of change, programme theory	Results framework
Adaptation	Aggregation, mixed methodology, data needs and priorities and dimensions		Data needs priorities and dimensions	Emergent
Contextual sensitivity	Development effectiveness, value for money, portfolio funding approaches		Value for money portfolio funding approaches	Sector development goals
Adaptation	Feedback loops, adaptive design, iterative frameworks, adaptive funding, measurement, reporting and learning		Adaptive design, iterative frameworks, funding, measurement, reporting and learning	Emergent
Relationality	Capability – leadership and managers		Capability leadership managers	Capacity in management
Contextual sensitivity	Data design approaches and data collection systems	Measurement	Data collection systems	Emergent
Relationality	Needs assessment – demand and needs		Demand and needs	Need analyses
Contextual sensitivity	Frameworks – outputs outcomes and impacts		Frameworks, outcomes and impacts	Results frameworks
Relationality	Mixed methodology for data sources, security and non-financials		Data sources, security and non-financials	Results frameworks



Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Adaptation	Frameworks, mixed methodology, indicators and targets		Indicators and targets	Results frameworks
Relationality	Mixed methodology – qualitative and quantitative data collection and analysis		Qualitative and quantitative data collection	Results frameworks
Adaptation	Feedback and learning		Learning and feedback	Emergent
Adaptation	Theory based evaluation, strategy and performance and M&E systems – role of M&E		Role of M&E	Emergent
Contextual sensitivity	Aligned frameworks, data and insights, reporting and accountability		Reporting and accountability	Emergent
Relationality	Interpersonal communications and functions – role and function of evaluation	Evaluation role and nature	Role and function	M&E processes
Adaptation	Theory based evaluation, nature and amount of monitoring and evaluation		Nature and amount monitoring and evaluation	M&E processes
Relationality	Integrated management functions and processes – structure and view of evaluation		Structure and view	M&E processes
Contextual sensitivity	Data aggregation and analysis, performance reporting		Performance reporting	M&E processes
Relationality	Capability and capacity in M&E – personal attributes including cultural competency	Evaluation practitioners' capacity and capability	Personal attributes	Capacity in M&E
Relationality	Capability and capacity in M&E – training and development including cultural understanding		Training and development	Capacity in M&E
Adaptation	Capability and capacity in M&E – skill and understanding of key business and evaluation concepts		Capacity and capability: skill and understanding of key business and evaluation concepts	Capacity in M&E
Adaptation	Devolution, leadership, integrated management functions and processes with outcomes and impacts focus, managerial capability, devolution and systematic embedded approach	Other agency capacity capability	Devolution, leadership and systematic embedded approach	Capacity in management



Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Relationality	Organisational culture, integrated systems for learning and performing, and evaluative capability		Organisational culture and system within organisation and capability	Emergent
<b>Institution contextual layer</b>				
Contextual, relationality	Strategy, design, inclusion, integrated evaluation design	Strategic management – strategy, design leadership, policy, learning, decision-making, culture, resources, communities	Strategy, design	Management processes
Relationality	Leadership, strategy, learning organisational culture, communication		Leadership	Management processes
Contextual sensitivity	Integrating planning and policy		Policy	Management processes
Adaptation	Integrating evaluative questions to strategy and performance, evidence based, decision-making		Decision-making	Management processes
Relationality, adaptation, contextual sensitivity	Organisational culture, interpersonal relations, learning, feedback loops, adaptive systems, inclusion		Organisational culture & learning	Emergent
Adaptation	Aligning investment with outcomes using data, resources		Resources	Priorities and budgets
Relationality, contextual sensitivity	Inclusion, service delivery, needs analysis, participatory, results measurement, regionalisation, decentralisation, communities of practice		Community inclusion	Emergent
Contextual sensitivity	Development theory and strategy, organisational strategy, system and sector approaches, policy, linkages, functions and structure	Performance management within institutions, audit, aggregation, reporting, learning	Development strategy and policy, linkages, functions and structure	Development strategy
Relational	Strategic alignment, systems, 'real world', business model as opposed to process model, integrated evaluation as a corporate function, sector system evaluation, management and performance management, complexity		Management and performance management	Management processes

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Adaptation	Auditing for both financial and non-financials		Audit	Management processes
Adaptation	Aggregation of results, analysis, interpretation, governance		Aggregation	Management processes
Adaptation	Aggregated reporting, feedback and learning, service delivery, interventions		Reporting and learning	Management processes
Relationality	Results-based approaches, partnership relationships, communication, motivation, knowledge sharing, inclusion, organisational culture	Interpersonal relationships communication, links, management administration processes, roles, IT systems	Communications and organisational culture	Emergent
Adaptation	Integrated functions – planning, strategy, evaluation, policy, finance, knowledge sharing, links and integration		Links and integration	Emergent
Relationality	Management administration and processes – human dynamics and relationships, systematic evaluative approach, organisational culture, communities of practice		Management administration and processes	Management processes
Adaptation	Management administration and processes – interpersonal roles and functions, integration and alignment of functions and roles, embedding evaluative approaches – outcomes and impacts focus, learning organisation		Interpersonal roles and functions	Emergent
Adaptation	IT systems, online processes and timeliness, data standards, regionalisation, integrated planning frameworks and reporting, data sharing, sustainability of service delivery and data collection, inclusion of providers and communities		IT systems	Emergent
Relationality	Interface with strategy, policy and M&E, inclusion, integrating strategy and evaluative approaches and planning, policy, business units, and M&E data collection	Monitoring and evaluation, research within institution	Interface with strategy, policy, and M&E	M&E processes
Adaptation	Evaluation demand and function – evaluation included as corporate function, design, feedback loops, internal evaluative capability and capacity within business teams		Evaluation demand and function	M&E processes
Contextual sensitivity	Reporting – levels of reporting strategic and operational, contribution analysis		Reporting	M&E processes
Relationality	Sector system evaluation, systems and systematic approaches, context relationships, data and evidence		Sector system evaluation	M&E processes

Higher level principles	Knowledge concept areas from qualitative research	Knowledge fields from research	Coded sub-codes	Original impact frame link
Contextual	Government planning – agency and Minister and Crown entities, constitutional approaches and processes, managerial theory and processes, strategic planning, policy, nature and role of evaluation	Role of ministers	Government planning-agency and Minister and Crown entities	Decision making
Adaptation	Reporting and accountability – reporting frequency, analyses, agency and sector performance and accountability, systematic data collections from regions		Reporting and accountability	Accountability
Adaptation	Sector approach, results frameworks, data collection, judgement, transparency	Value for money	Approach	Accountability
Adaptation	Development outcomes and adaptive inputs, information and data required, theory of change, results frameworks, systematic data collection – bottom-up, programme, thematic and strategic, data aggregation and analysis, sustainability and needs analysis, feedback loops		Development outcomes and adaptive inputs information and data required	Emergent

## Appendix G: Full summary table of identified concepts from research

Table 18. Concepts identified from the research

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
<b>Layer 3: Wider infrastructure contextual layer</b>		
Impact model dimensions: Improved lives for people and the environment; Improved development effectiveness; Improved aid effectiveness.		
Development – approaches, power & democracy, culture, national capacity		
	<ul style="list-style-type: none"> <li>• Development theory and approach, context, wellbeing, government role and community inclusion</li> <li>• Paris Declaration principles, centralisation and devolution, power and control, mutual accountability</li> <li>• Development results – focus on outcomes, statistics and data</li> </ul>	
Governance		
	<ul style="list-style-type: none"> <li>• Development principles and approaches with inclusion, partnership, and participation</li> <li>• Constitutional arrangements and managerialism</li> <li>• Centralisation and devolution, mutual accountability, agency maturity</li> <li>• Government and management accountability and democracy and inclusion, accountability frameworks and people performance,</li> </ul>	<ul style="list-style-type: none"> <li>• Service delivery, community inclusion and evaluative feedback, iterative adaptation, and results</li> </ul>

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
	<ul style="list-style-type: none"> <li>• Sector and country portfolio planning and investing, aligning financial systems including NGOs private sector.</li> <li>• Explicit theories of change, prioritising outcomes, and investment; politics, donor alignment.</li> <li>• Systematic planning and prioritisation approaches and standardisation, systems functions, flexibility, regionalisation, embedding evaluative approaches.</li> </ul>	
Central agencies		
	<ul style="list-style-type: none"> <li>• Role of central agencies – development, evaluation, and management: State sector model, explicit function for evaluation, audit role, central agency role with performance and accountability.</li> <li>• Agency view of central agencies – extended approaches and processes to include the use of results and outcomes frameworks, accountability, central agency role, donors aligning to country priorities.</li> </ul>	
Accountability – performance management, audit, evaluation, financial, country systems		
<ul style="list-style-type: none"> <li>• Audit guidance and standards for non-financial information, centralisation and devolution, mutual accountability, agency maturity (A/NZ)</li> </ul>	<ul style="list-style-type: none"> <li>• Accountability and alignment, values and value of outcomes, results and outcomes frameworks, inclusion, integration, performance data, resourcing, accountability, devolution, non-financial information, sector approaches</li> <li>• Performance management systems, state sector model, managing for outcomes</li> </ul>	<ul style="list-style-type: none"> <li>• Country development, results and role of monitoring &amp; evaluation, country-level requirements for results and accountability, activity level requirements with alignment, investment approach, evaluation role and approaches</li> </ul>

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
<ul style="list-style-type: none"> <li>Provincial mandate for coordination, administration, service delivery and M&amp;E (PNG)</li> </ul>	<p>theory, mandate for evaluation, organisational culture for sector and institutional performance and learning, coproduction</p> <ul style="list-style-type: none"> <li>Results frameworks and evaluation, integration, systematic and flexible approaches, M&amp;E focus, learning organisation, alignment of country systems and inputs, stakeholder engagement</li> <li>Financial funding for outputs and outcomes, investment, prioritisation and allocation, portfolio budgeting, leadership, manager capability, country, sector, and thematic alignment</li> </ul>	<ul style="list-style-type: none"> <li>Regionalisation, development planning and alignment of investment, incrementality, evaluative discussions, community involvement</li> </ul>
Donors – govt to govt cooperation, role of monitoring & evaluation		
	<ul style="list-style-type: none"> <li>Government role and government to government cooperation, applicability to other country systems, context, constitutions, relationships, aligning to country systems</li> </ul>	<ul style="list-style-type: none"> <li>Country development, results, and role of monitoring &amp; evaluation. Country-level requirements for results measurement and accountability</li> <li>Activity-level requirements with alignment and investment approach. Evaluation role and approaches at activity-level.</li> </ul>

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
<b>Layer 1: Ideas, interventions and frameworks layer (Repositioned as Strategic Evaluation)</b>		
Impact model dimensions: Increased use of results frameworks, country systems, ownership, sector development goals, need analyses, capacity in management, capacity in M&E.		
Architecture of ideas, interventions, and frameworks		
<ul style="list-style-type: none"> <li>Theory based evaluation, programme theory, theories of change, outcomes, contribution, attribution</li> </ul>	<ul style="list-style-type: none"> <li>Devolution, leadership, programme theory – change of state, theory of change, context, processes, and outcomes</li> <li>Strategic planning and emergent strategic intent, theory-based evaluation, portfolio-based funding</li> </ul>	<ul style="list-style-type: none"> <li>Values – inclusion, participatory, collaboration and consultation</li> </ul>
Subgroups: ethnicity, gender, and regions		
	<ul style="list-style-type: none"> <li>Inclusion, participatory, segmentation</li> <li>Theory based evaluation, subgroups, indicators and targets, segmentation, and reporting</li> </ul>	<ul style="list-style-type: none"> <li>Regionalisation, decentralisation</li> </ul>
Use of frameworks		
<ul style="list-style-type: none"> <li>Theory based evaluation, strategic planning, strategic alignment, priorities and frameworks-layers, theory of change, programme theory</li> </ul>	<ul style="list-style-type: none"> <li>Context, country and sector systems, adaptive and emergent strategy, sector, and system driven and institutional performance</li> <li>Development effectiveness, value for money, portfolio funding approaches</li> <li>Capability – leadership and managers</li> </ul>	<ul style="list-style-type: none"> <li>Values, prioritisation, inclusion, interpersonal dialogue on key priorities, contextual inclusive leadership</li> <li>Aggregation, mixed methodology, data needs and priorities and dimensions</li> </ul>

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
		<ul style="list-style-type: none"> <li>Feedback loops, adaptive design, iterative frameworks, adaptive funding, measurement, reporting and learning</li> </ul>
Measurement		
<ul style="list-style-type: none"> <li>Frameworks – outputs outcomes and impacts, indicators and targets</li> <li>Mixed methodology for data sources, security, and non-financials</li> <li>Mixed methodology – qualitative and quantitative data collection and analysis</li> </ul>	<ul style="list-style-type: none"> <li>Needs assessment – demand and needs</li> </ul>	<ul style="list-style-type: none"> <li>Data design approaches and data collection systems.</li> <li>Feedback and learning</li> <li>Theory based evaluation, strategy and performance and M&amp;E systems -role of M&amp;E</li> <li>Aligned frameworks, data and insights, reporting and accountability</li> </ul>
Management capability		
	<ul style="list-style-type: none"> <li>Devolution, leadership, integrated management functions and processes with outcomes and impacts focus, managerial capability, devolution and systematic embedded approach</li> </ul>	<ul style="list-style-type: none"> <li>Organisational culture, integrated systems for learning and performing, and evaluative capability</li> </ul>
Capability and capacity in M&E		
	<ul style="list-style-type: none"> <li>Personal attributes including cultural competency</li> <li>Training and development including cultural understanding</li> <li>Skills and understanding of key business and evaluation concepts</li> </ul>	
Evaluation role and nature		
	<ul style="list-style-type: none"> <li>Communications and functions – role and function of evaluation</li> </ul>	



Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
	<ul style="list-style-type: none"> <li>• Theory based evaluation, nature and amount of monitoring and evaluation</li> <li>• Integrated management functions and processes-structure and view of evaluation</li> <li>• Data aggregation and analysis, performance reporting.</li> </ul>	
<b>Layer 2: Institution contextual layer</b>		
Impact model dimensions: Development strategy, management processes, priorities and budgets, M&E processes, improved decision-making, increased accountability.		
Strategic management – strategy, design leadership, policy, learning, decision-making, culture, resources, communities		
	<ul style="list-style-type: none"> <li>• Strategy, design, inclusion, integrated evaluation design</li> <li>• Leadership, strategy, learning organisational culture, communication</li> <li>• Integrating planning and policy</li> <li>• Integrating evaluative questions to strategy and performance, evidence based, decision-making</li> <li>• Aligning investment with outcomes using data, resources.</li> </ul>	<ul style="list-style-type: none"> <li>• Organisational culture, interpersonal relations, learning, feedback loops, adaptive systems.</li> <li>• Community inclusion, service delivery, needs analysis, participatory, results measurement</li> <li>• Regionalisation, decentralisation, communities of practice.</li> </ul>
Interpersonal relationships, communication, links, management administration processes, roles		
	<ul style="list-style-type: none"> <li>• Management administration and processes – human dynamics and relationships, systematic evaluative approach, organisational culture, communities of practice.</li> </ul>	<ul style="list-style-type: none"> <li>• Communications and organisational culture: Results-based approaches, partnership relationships, communication, motivation, knowledge sharing, inclusion, organisational culture.</li> </ul>

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
		<ul style="list-style-type: none"> <li>• Integrated functions – planning, strategy, evaluation, policy, finance, knowledge sharing, links, and integration.</li> <li>• Management administration and processes Interpersonal roles and functions, integration and alignment of functions and roles, embedding evaluative approaches – outcomes and impacts focus, learning organisation.</li> <li>• IT systems, online processes and timeliness, data standards, regionalisation, integrated planning frameworks and reporting, data sharing, sustainability of service delivery and data collection, inclusion of providers and communities.</li> </ul>
Value for money		
	<ul style="list-style-type: none"> <li>• Sector approach, results frameworks, data collection, judgement, transparency.</li> </ul>	<ul style="list-style-type: none"> <li>• Development outcomes and adaptive inputs, information and data required, theory of change, results frameworks, systematic data collection – bottom-up</li> <li>• Programme, thematic and strategic, data aggregation and analysis, sustainability and needs analysis, feedback loops.</li> </ul>
Role of Ministers and State		

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
	<ul style="list-style-type: none"> <li>Government planning – agency and Minister and Crown entities, constitutional approaches and processes, managerial theory and processes, strategic planning, policy, nature and role of evaluation</li> <li>Reporting and accountability – reporting frequency, analyses, agency and sector performance and accountability, systematic data collections from regions.</li> </ul>	
Performance management within institutions, audit, aggregation, reporting, learning		
<ul style="list-style-type: none"> <li>Auditing for both financial and non-financials.</li> </ul>	<ul style="list-style-type: none"> <li>Strategic alignment, systems, ‘real world’, business model as opposed to process model, integrated evaluation as a corporate function, sector system evaluation, management and performance management, and complexity.</li> <li>Aggregation of results, analysis, interpretation, governance.</li> <li>Reporting and learning: Aggregated reporting, feedback and learning, service delivery, interventions.</li> </ul>	
Monitoring and evaluation, research within institution		
	<ul style="list-style-type: none"> <li>Interface with strategy, policy and monitoring and evaluation – inclusion, integrating strategy and evaluative approaches and planning, policy, business units, and M&amp;E data collection</li> <li>Evaluation demand and function – evaluation included as a corporate</li> </ul>	

Concepts showing progress	Concepts requiring more attention	Concepts requiring inclusion
	<p>function, design, feedback loops, internal evaluative capability and capacity within business teams</p> <ul style="list-style-type: none"> <li>• Levels of reporting – strategic and operational, contribution analysis</li> <li>• Sector system evaluation, systems and systematic approaches, context, relationships, data and evidence.</li> </ul>	

## Appendix H: List of organisations who participated in this research

Aotearoa New Zealand	Papua New Guinea
<b>Central agencies:</b>	<b>Central agencies:</b>
Audit New Zealand	Department of National Planning and Monitoring
Office of Auditor General	Department of Treasury
State Services Commission	Department of Finance
The Treasury	Vision 2050
<b>Cultural or regional agencies</b>	<b>Cultural or regional agencies</b>
Te Puni Kōkiri (Maori Development Agency)	Department of Provincial and Local Government Affairs
Ministry of Pacific Peoples	
<b>Line agencies:</b>	<b>Line agencies:</b>
Ministry of Transport	Department of Works
Ministry of Education	Department of Community Development
Ministry of Economic Development	Department of Health
Ministry of Culture and Heritage	Department of Education
Ministry for the Environment	Department Agriculture
Ministry of Agriculture and Forestry	Justice Sector

Aotearoa New Zealand	Papua New Guinea
Ministry of Health	Transport Sector
Ministry of Justice	<b>Development partners:</b>
Ministry of Social Development	New Zealand Ministry of Foreign Affairs and Trade (includes former NZAID)
	Australian Department of Foreign Affairs and Trade (includes former AusAID)
	The World Bank
	Asian Development Bank

