

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.

Developing and Testing a Performance Measurement Framework for the Australasian Nonprofit Healthcare Sector

A thesis presented in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
in
Engineering
at Massey University, Manawatu,
New Zealand.

H. Ishani Buddika Soysa

2017

Abstract

Irrespective of the sector to which an organisation belongs, measuring and monitoring the strategic performance is of paramount importance for survival and growth of the organisation. While the balanced scorecard (BSC) remains the most widely recognised performance measurement (PM) model by the academia and the practitioners, the theoretical validity and generalisability of it for the nonprofit sector remains to be tested via samples. In addition, the technical requirement of *strategy maps* acts as a deterrent to the uptake of the BSC in the nonprofit sector.

In this research, the researcher theorised and empirically validated the nonprofit version of the BSC (nonprofit BSC) using data collected from Australasian healthcare NPOs, using an exploratory sequential mixed methods research platform. First, a conceptual model was developed based on the literature on the BSC and relevant other PM models. Next, the conceptual model was advanced into a causal predictive model (referred to as the main theoretical model in this thesis) using qualitative data collected from senior managers belonging to nine Australasian healthcare NPOs. In keeping the nonprofit BSC, the researcher's main theoretical model posits that the Mission of the NPO drives its Strategy, which together drives its Financial Health, People, and the Infrastructure to enable the Processes to deliver mission-related outcomes; the mission-related outcomes being represented by Client Satisfaction, People Satisfaction, and Donor Satisfaction. Finally, quantitative data collected from Australian and New Zealand healthcare NPOs, in the form of survey responses ($n = 223$), were used to test the researcher's theoretical models, using the *partial least squares based structural equation modelling* approach. The theoretical and practical contributions of the study are: (a) development of a theoretical model that underpins the nonprofit BSC; (b) operationalisation of the constructs of the model, which in effect, operationalises the nonprofit BSC; and (c) empirical development of an overall organisational performance index linked to an associated scoring system. The theoretical and practical implications of the study (Chapters Six and Eight) include how the main theoretical model can be used to predict and explain successful performance improvement interventions and unsuccessful performance improvement interventions, how an organisation could develop its own performance measures without having to use a strategy map, and general guidelines for performance improvement including how the scoring system can be used for self-assessment. The limitations and future research directions have been outlined in the conclusions chapter.

Acknowledgements

A doctoral study is a journey with ups and downs, and mine is no different! Many people assisted me in my journey in various ways. I hope I have not omitted the names of all the key people who assisted me. First and foremost, I am indebted to my supervisor, Dr Nihal Jayamaha, for his invaluable guidance throughout my research. His rigour, patience, interest and endless encouragement have given me strength to keep going, every step of the way. I am also grateful to my co-supervisor, Professor Nigel Grigg, for his contagious enthusiasm and his confidence that I could reach excellence, if I really tried. I also acknowledge that my supervisors were the second and third authors of all my publications.

I am also grateful to my informants—the nine senior managers—who provided me written and verbal responses in my qualitative study. Thank you very much for your valuable time and patience. I am also thankful to the other senior managers who responded to my quantitative survey. I am also indebted to the anonymous reviewers who reviewed my journal articles and conference papers. Your feedback was insightful and that helped me to improve my work. I also thank the conference attendees who provided me verbal feedback on my presentations.

I also owe a big thank to Massey University for granting me the Massey University Doctoral Scholarship, which greatly reduced my financial burden. I also thank the administration staff of the School of Engineering and Advanced Technology (SEAT), Massey University, particularly Michele Wagner, Linda Lowe, Lateia Cornwall, and Dilantha Punchihewa and the SEAT Financial Department, the Scholarship Office, the Graduate Research School, the International Office, and the Librarians for their support throughout my study.

Last, but certainly not the least, a special thanks goes to my family: to my parents, who raised me and gave me the ambition to pursue new challenges; and to my great husband and my lovely daughter, for their immense support.

Ishani Soysa
25 July 2016

Table of Contents

ABSTRACT.....	I
ACKNOWLEDGEMENTS.....	II
TABLE OF CONTENTS.....	III
LIST OF ACRONYMS.....	XI
LIST OF FIGURES	XII
LIST OF TABLES.....	XV
CHAPTER ONE:INTRODUCTION	1
1.1. INTRODUCTION.....	1
1.2. BACKGROUND TO THE RESEARCH	2
1.2.1. The Nature, Diversity, and Growth of the Nonprofit Sector	3
1.2.2. Performance Measurement and Performance Measurement Frameworks	7
1.3. RESEARCH MOTIVATION	10
1.4. AIMS AND OBJECTIVES OF THE STUDY	11
1.4.1. Research Aim.....	11
1.4.2. Research Objectives.....	11
1.5. RESEARCH QUESTIONS	12
1.6. RESEARCH CONTRIBUTION	14
1.7. LIMITATIONS OF THE STUDY	15
1.8. THE STRUCTURE OF THE THESIS	16
CHAPTER TWO: INTRODUCTION TO NONPROFIT ORGANISATIONS	21
2.1. INTRODUCTION.....	21
2.2. NONPROFIT ORGANISATIONS: THE GENESIS, DEFINITION AND ECONOMIC THEORIES	22
2.2.1. The Genesis of NPOs.....	22
2.2.2. Defining NPOs.....	23
2.2.3. The International Classification of NPOs.....	24
2.2.4. Theories That Explain the Existence of NPOs	26
2.3. GOVERNANCE AND ACCOUNTABILITY IN NPOS.....	27
2.3.1. Nonprofit Governance	27
2.3.1.1. Agency Theory.....	28
2.3.1.2. Stakeholder Theory	29
2.3.1.3. Resource Dependence Theory.....	30

2.3.2. Accountability.....	30
2.4. NONPROFIT SECTOR IN AUSTRALASIA.....	32
2.5. CLASSIFICATION OF HEALTHCARE ORGANISATIONS	34
2.5.1. Classification of Australasian Nonprofit Healthcare Organisations.....	35
2.6. PERFORMANCE MEASUREMENT CHALLENGES IN NPOS	37
2.7. CHAPTER CONCLUSION	38
CHAPTER THREE: CORE LITERATURE ON PERFORMANCE MEASUREMENT SYSTEMS... 39	
3.1. INTRODUCTION.....	39
3.2. THE RATIONALE FOR PERFORMANCE MEASUREMENT AND ITS EVOLUTION	40
3.3. PERFORMANCE MEASUREMENT FRAMEWORKS	44
3.3.1. Role of PM Frameworks and Empirical Support for Their Effectiveness.....	45
3.3.2. The Conceptual Foundation of the BSC	48
3.3.3. The Conceptual Foundation of the PP	56
3.3.3.1. PP Vs BSC: Researcher’s Synthesis	58
3.3.4. The Conceptual Foundation of the Baldrige Excellence Framework and the EFQM Excellence Model	59
3.3.4.1. Baldrige Excellence Framework.....	59
3.3.4.2. EFQM Excellence Model.....	62
3.4. PERFORMANCE MEASUREMENT FRAMEWORKS FOR NPOS	64
3.4.1. The Balanced Scorecard for NPOs	64
3.4.2. The Performance Prism and Quality Award/Excellence Frameworks	68
3.4.3. The Production of Welfare Framework.....	69
3.4.4. The Logical Framework Approach/Logic Model.....	70
3.5. PRIOR STUDIES THAT EXAMINED DIFFERENT FACETS OF VALIDITY OF THE BSC 71	
3.5.1. BSC Implementation Studies.....	71
3.5.2. Conceptual Papers	74
3.5.3. Empirical Studies.....	75
3.6. KNOWLEDGE GAPS AND RESEARCH QUESTIONS	77
3.7. CHAPTER CONCLUSION	80
CHAPTER FOUR: THE RESEARCH METHODOLOGY	81
4.1. INTRODUCTION.....	81
4.2. RESEARCH PARADIGMS	82
4.2.1. Positivism and Postpositivism	82
4.2.2. Constructivism.....	83

4.2.3. Participatory/Advocacy Worldview	85
4.2.4. Pragmatism	85
4.3. MIXED METHODS DESIGNS	86
4.3.1. Introducing data collection phases.....	86
4.3.2. Classifying the specific mixed methods design adopted	91
4.3.3. Justification of the Researcher’s Paradigm.....	95
4.4. DEVELOPING THE CONCEPTUAL MODEL	96
4.5. QUALITATIVE DATA COLLECTION: REFINING THE CONCEPTUAL MODEL	101
4.5.1. Sample Selection for the Semi-Structured Questionnaire	102
4.5.2. Designing the Semi-Structured Questionnaire	105
4.5.3. The Follow-up Interview	107
4.5.4. Deriving the Final Model.....	108
4.5.5. Operationalising the Constructs of the Final Model.....	110
4.6. QUANTITATIVE DATA COLLECTION	111
4.6.1. Determining the Sample Size	111
4.6.2. Developing the Survey Questionnaire	112
4.6.3. Selecting the Sampling Frame and Full Scale Data Collection	114
4.6.4. Tests Conducted Prior to Full Scale Data Collection	116
4.6.4.1. Content validity and Face Validity Assessment.....	116
4.6.4.2. Pre-testing and Pilot Testing	117
4.7. QUANTITATIVE DATA ANALYSIS.....	118
4.7.1. Data Screening.....	118
4.7.2. Verifying the Absence of Common Method Bias	119
4.7.3. Establishing Scale Reliability	119
4.7.4. Establishing Construct Validity	120
4.7.5. Model Testing	121
4.8. DEVELOPING AND TESTING THE ANCILLARY MODEL	121
4.9. JUSTIFYING THE QUALITY OF THE RESEARCH.....	123
4.9.1. Planning Quality	125
4.9.2. Design Quality	125
4.9.3. Data Quality.....	128
4.9.4. Interpretive Rigour.....	131
4.9.5. Inference Transferability	133
4.9.6. Other Considerations	134

4.9.6.1. Samples, Sampling, and the Sample Frames	134
4.9.6.2. Ethical Considerations	136
4.10. CHAPTER SUMMARY AND CONCLUSION	136
CHAPTER FIVE: REVIEW OF POSITIVISTIC TECHNIQUES USED IN TESTING	
THEORETICAL MODELS.....	138
5.1. INTRODUCTION.....	138
5.2. VALIDITY OF A THEORETICAL MODEL FROM A QUANTITATIVE RESEARCHER’S	
STANDPOINT	138
5.2.1. The Building Blocks of a Theoretical Model	138
5.2.2. Modelling of Constructs	140
5.2.2.1. The Standard Form of Modelling a Construct	140
5.2.2.2. Modelling Formative Constructs	143
5.2.2.3. Modelling Hierarchical Constructs	144
5.3. FORMS OF VALIDITY.....	146
5.3.1. Content Validity.....	146
5.3.2. Predictive validity	146
5.3.3. Construct Validity.....	147
5.3.3.1. Reliability and Internal Consistency	147
5.3.3.2. Methods of Establishing Construct Validity	148
5.3.4. Statistical conclusion Validity	149
5.3.5. Internal and External Validity	149
5.3.5.1. Internal Validity	149
5.3.5.2. External Validity	150
5.4. STRUCTURAL EQUATION MODELLING	151
5.4.1. The Covariance Approach to SEM.....	151
5.4.2. The Partial Least Squares Approach to SEM	152
5.5. VALIDITY CONSIDERATIONS IN CASE STUDY RESEARCH.....	154
5.6. CHAPTER CONCLUSION	155
CHAPTER SIX: QUALITATIVE DATA ANALYSIS AND RESULTS	156
6.1. INTRODUCTION.....	156
6.2. SUMMARY DATA ON CASE STUDY ORGANISATIONS	157
6.2.1. The Case study Organisations	157
6.3. OUTCOMES OF THE SEMI-STRUCTURED QUESTIONNAIRE AND THE FOLLOW-UP	
INTERVIEWS	161

6.4. DISCUSSION OF KEY THEMES AND SUBTHEMES	168
6.4.1. Mission and Core Values	168
6.4.2. Strategy	169
6.4.3. Organisational Capabilities, and People Development	170
6.4.4. Economic Sustainability	173
6.4.5. Processes	175
6.4.6. Stakeholder Satisfaction	177
6.5. PRESENTING THE OPERATIONAL DESCRIPTIONS OF THE UNDERLYING PM DIMENSIONS.....	180
6.6. DEVELOPMENT OF THE SURVEY ITEMS OF THE QUANTITATIVE QUESTIONNAIRE/MEASUREMENT INSTRUMENT	184
6.6.1. Questionnaire Items that Capture Mission	184
6.6.2. Questionnaire Items that Capture Strategy	185
6.6.3. Questionnaire Items that Capture People	185
6.6.4. Questionnaire Items that Capture Organisational Infrastructure	185
6.6.5. Questionnaire Items that Capture Financial Health	185
6.6.6. Questionnaire Items that Capture Processes	186
6.6.7. Questionnaire Items that Capture Client Satisfaction	186
6.6.8. Questionnaire Items that Capture People Satisfaction	187
6.6.9. Questionnaire Items that Capture Donor Satisfaction	187
6.7. THEORETICAL AND PRACTICAL IMPLICATIONS	187
6.7.1. Theoretical Implications of the Case Study Findings	187
6.7.2. Practical Implications of the Case Study Findings	189
6.8. CHAPTER CONCLUSION	194
CHAPTER SEVEN: DESCRIPTIVE STATISTICS.....	196
7.1. INTRODUCTION.....	196
7.2. STATISTICS OF THE QUANTITATIVE STUDY	196
7.2.1. Response Pattern During the Survey Period.....	197
7.2.2. The Profile of the Survey Participants	198
7.2.3. Sector Representation	198
7.2.4. The Participants' Familiarity with Performance Measurement Systems	199
7.2.5. The Chronological Age of the Responds' Organisation.....	200
7.2.6. Funding Streams	201
7.2.7. Distribution of Paid Employees	202

7.3. ANALYSIS OF INDIVIDUAL SURVEY ITEMS	203
7.3.1. Individual Survey Item Scores.....	204
7.3.2. Examination of Bivariate Correlations between Questionnaire Items	212
7.3.3. Aggregation of Questionnaire Item Scores to Examine Data at Performance Dimension/Construct Level.....	215
7.4. CHAPTER CONCLUSION	217
CHAPTER EIGHT: EMPIRICAL TEST RESULTS AND DISCUSSION OF THE THEORETICAL MODELS.....	218
8.1. INTRODUCTION.....	218
8.2. RESULTS FOR MEASUREMENT VALIDITY	220
8.2.1. Data Screening.....	220
8.2.2. Testing for Absence of Common Method Bias	221
8.2.3. Scale Reliability.....	221
8.2.4. Construct Validity.....	223
8.2.4.1. Assessing the construct validity of the formative construct	226
8.3. TESTING THE HYPOTHESISED THEORETICAL RELATIONSHIPS	227
8.3.1. The Overall Goodness of Fit of the Model.....	228
8.4. DISCUSSING THE RESULTS OF THE MAIN MODEL FROM A THEORETICAL PERSPECTIVE.....	233
8.5. DISCUSSING THE RESULTS ON THE MAIN MODEL FROM A PRACTICAL PERSPECTIVE.....	235
8.5.1. Concluding Discussion on the Third Research Objective	238
8.6. ANALYSING THE COUNTRY EFFECT ON THE HYPOTHESISED RELATIONSHIPS... 238	
8.6.1. Results on the Australian Model	240
8.6.2. Results on the New Zealand Model.....	243
8.6.3. Testing the Differences in Path Coefficients - the Australian Model versus the New Zealand Model.....	245
8.7. RESULTS AND DISCUSSION OF THE EMPIRICAL WEIGHTS FOR THE PERFORMANCE DIMENSIONS.....	250
8.7.1. Discussion of Empirical Results on the Ancillary Model	253
8.8. IMPLICATIONS FOR MANAGEMENT PRACTICE AND PERFORMANCE MEASUREMENT.....	256
8.8.1. Operationalisation of the Constructs	256
8.8.2. Predicting and Explaining how Performance can be Improved	258

8.8.3. Implication of the Ancillary Model	260
8.9. COMPARING THE RESEARCHER’S PM MODEL WITH THE NONPROFIT BSC AND ORGANISATIONAL EXCELLENCE MODELS	263
8.10. CHAPTER CONCLUSION	267
CHAPTER NINE: CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH	270
9.1. INTRODUCTION.....	270
9.2. KNOWLEDGE GAPS, RESEARCH QUESTIONS, RESEARCH OBJECTIVES, AND THE MIXED METHODS APPROACH	271
9.2.1. The Basis for the Research Questions and Objectives	271
9.2.2. The Mixed Methods Approach	273
9.3. FINDINGS FOR EACH RESEARCH OBJECTIVE.....	276
9.3.1. Findings on Objective 1	276
9.3.2. Findings on Objective 2.....	277
9.3.3. Findings on Objective 3	280
9.3.4. Findings on Objective 4.....	283
9.3.5. Findings on Objective 5.....	285
9.3.6. Findings on the General Objective	288
9.4. JUSTIFYING THE ORIGINAL CONTRIBUTION	289
9.4.1. The Theoretical Contribution	289
9.4.1.1. Theory Development.....	289
9.4.1.2. Theory Testing	290
9.4.2. The Practical Contribution.....	291
9.5 METHODOLOGICAL LESSIONS LEARNT.....	291
9.6. LIMITATIONS AND FURTHER RESEARCH DIRECTIONS	293
9.6.1. Enhancing the Generalisability	294
9.6.2. Studying Potential Moderating Effects.....	294
9.6.3. A Longitudinal Study for Enhancing the Practical Usability	294
9.6.4. Adapting to the Public Sector	295
REFERENCES.....	296
APPENDICES	318
APPENDIX A: The Semi-Structured Questionnaire Used in the Case Study.....	319
APPENDIX B: Invitation (Email) to Participants: the Quantitative Questionnaire. ...	329
APPENDIX C: The Online Survey Questionnaire.....	331
APPENDIX D: Massey University Human Ethics Approval – Low Risk.....	339

APPENDIX E: Quality Assessment of the Ancillary Model.....340

APPENDIX F: The Author’s Publications Related to her Doctoral Research: the DRC
16 Forms..... 343

APPENDIX G: Researcher’s Journal Articles [Page 1 only]..... 350

List of Acronyms

ACNC	Australian Charities and Not-for-profits Commission
AMOS	Analysis of Moment Structures
ASQ	American Society of Quality
AVE	Average Variance Extracted
BSC	Balanced Scorecard
BEF	Baldrige Excellence Framework
CBSEM	Covariance Based Structural Equation Modelling
CEO	Chief Executive Officer
CFA	Confirmatory Factor Analysis
CR	Composite Reliability
CRM	Cause-Related Marketing
CSF	Critical Success Factors
EFQM	European Foundation of Quality Management
EFA	Exploratory Factor Analysis
EEM	EFQM Excellence Model
GDP	Gross Domestic Products
GM	General Manager
GVA	Gross Value Added
ICNPO	International Classification of Nonprofit Organisations
ICT	Information and Communications Technology
KPI	Key Performance Indicator
LFA	Logical Framework Approach
NIST	National Institute of Science and Technology
NPO	Nonprofit Organisation
OPI	Overall Performance Index
PCA	Principal Components Analysis
PLSBSEM	Partial Least Squares Based Structural Equation Modelling
PM	Performance Measurement
PP	Performance Prism
ROI	Return on Investment
SEM	Structural Equation Modelling

LIST OF FIGURES

Figure 1.1: Nonprofit sector and GDP growth rates of five selected countries	6
Figure 1.2: The nexus between research questions and research objectives	13
Figure 1.3 : The structure of the thesis	20
Figure 3.1: Number of publications on four selected PM frameworks.....	45
Figure 3.2: The original BSC framework published by Kaplan and Norton	49
Figure 3.3: The second generation BSC framework.....	50
Figure 3.4: A typical strategy map of a for-profit organisation embodied in a third generation BSC	51
Figure 3.5: The PP framework and an example of a success map.....	56
Figure 3.6: The Baldrige excellence framework.....	60
Figure 3.7: The EFQM excellence model.....	63
Figure 3.8: The nonprofit BSC framework.....	64
Figure 3.9: A strategy map of a NPO	66
Figure 3.10: The hierarchical arrangement of the BSC perspectives for a NPO.....	67
Figure 3.11: Production of welfare framework.....	69
Figure 3.12: The basic Logic Model.....	70
Figure 4.1: The strategy implemented to answer the research questions.....	87
Figure 4.2: Flow of the overall research	89
Figure 4.3: The flow of researcher’s exploratory sequential design.....	94
Figure 4.4: The conceptual model developed for Australasian nonprofit healthcare organisations	98
Figure 4.5: The flow chart for sample size determination for qualitative data collection	104
Figure 4.6: The finalised model - the main theoretical model.....	109

Figure 4.7: The hierarchical model on overall performance.....	123
Figure 4.8: The customised quality framework used by the researcher.....	124
Figure 5.1: Specifying the measurement model and the structural model in graphical format	142
Figure 5.2: Illustration of the measurement model of a formative construct.....	144
Figure 5.3: Alternative forms of representing a second-order construct	145
Figure 6.1: The case study organisations by subcategory.....	157
Figure 6.2: Classification of organisations by age.....	158
Figure 6.3: A screen dump of an output generated by Interview Streamliner software	162
Figure 6.4: The final theoretical model.....	183
Figure 7.1: Survey representation by country	196
Figure 7.2: Response rate during the survey.....	197
Figure 7.3: Participants by designation.....	198
Figure 7.4: Health subcategory wise responses	199
Figure 7.5: Use of a PM system.....	199
Figure 7.6: Participants' familiarity with PM systems.....	199
Figure 7.7: Age of the organisation	201
Figure 7.8: Level of government funding	202
Figure 7.9: Distribution of the paid employees.....	203
Figure 7.10: Mean scores of survey items in the chronological order	205
Figure 7.11: Mean scores of survey items based on the participants' familiarity of PM	210
Figure 7.12: Mean scores of survey items by country for participants who rated themselves as being very much familiar with PM systems	211

Figure 7.13: Mean scores of survey items by country for participants who rates themselves as being somewhat familiar with PM systems	211
Figure 7.14: Mean scores of survey items by country for participants who rated themselves as being very little familiar with PM systems	212
Figure 7.15: Mean scores of survey items based aggregated at performance dimension level.....	215
Figure 7.16: Mean scores of survey items aggregated at PM dimension level and by country	216
Figure 8.1: Estimated structural model parameters and the significance of hypothesised relationships	228
Figure 8.2: Parameter estimates in the Australian model	242
Figure 8.3: Parameter estimates in the New Zealand model	245
Figure 8.4: The hierarchical model on overall performance.....	251
Figure 8.5: Research questions answered and research objectives achieved.....	269
Figure 9.1: The links between the research questions and research objectives as well as the links between the thesis chapters	275
Figure 9.2: The finalised theoretical model on strategic performance	276
Figure 9.3: The structural relationships between constructs and parameter estimates	282
Figure 9.4: The hierarchical model on overall performance.....	284

LIST OF TABLES

Table 2.1: International Classification of Nonprofit Organizations.....	24
Table 2.2: The ICPNO Sub-classification of the Health Sector.....	35
Table 2.3: Classification of the Australian Nonprofit Healthcare Sector	36
Table 2.4: Classification of the New Zealand Nonprofit Healthcare Sector.....	36
Table 4.1: The Profile of the NPOs and the Respondents.....	105
Table 4.2: The Content of the Semi-Structured Questionnaire in Summery Form.....	105
Table 5.1: A Systematic Model Evaluation Procedure for PLSBSEM.....	153
Table 6.1: Profile of the Participants and Their Organisations.....	159
Table 6.2: Questionnaire Numbers Corresponding to Keywords	162
Table 6.3: Themes, Subthemes and Related Information	163
Table 6.4: The Operational Descriptions of Each PM Dimension.....	180
Table 6.5: A Specimen Translation to the Questionnaire Format.....	184
Table 6.6 : The Critical Success Factors and Suggested KPIs.....	190
Table 7.1: Measures of Central Tendency and the Dispersion	205
Table 7.2: The Mean Scores Based on Country.....	208
Table 7.3: Bivariate Correlations between Survey Items.....	214
Table 8.1: The Reliability Statistics for Conventional Constructs.....	222
Table 8.2: Latent Variable Correlations for Discriminant Validity	224
Table 8.3: Loadings and Cross-Loadings Based on the Reflective Model	225
Table 8.4: Collinearity Assessment.....	229
Table 8.5: Predictive Relevance of Endogenous Reflective Constructs	231
Table 8.6: The Effect Sizes of the Predictors of the Relevant Endogenous Constructs	232
Table 8.7: Reliability Statistics of the Australian Model	240

Table 8.8: Loading and Cross-Loading Information for the Australian Model to Examine Construct Validity.....	241
Table 8.9: Reliability Statistics of the New Zealand Model	243
Table 8.10: Loading and Cross-Loading Information for the New Zealand Model to Examine Construct Validity.....	244
Table 8.11: The Results on the Hypotheses	246
Table 8.12: Parameters Associated With Unsupported Hypotheses	248
Table 8.13: Standardised Regression Coefficients of Constructs and Items	252
Table 8.14: The Summary Description of the Point Scoring System	252
Table 9.1: Operational Descriptions of Each PM Dimension.....	278
Table 9.2: The Point Scoring System.....	284

CHAPTER ONE

INTRODUCTION

We cannot solve our problems with the same thinking we used when we created them.

Albert Einstein

1.1. INTRODUCTION

One thing that sets humans apart from other species in the planet is their quest for improvement! Humans would like to monitor the performance of what they do (both individually and collectively) now, so that they could create new goals for the future to improve their current performance. An Organisation is a collection of individuals working together to achieve a common goal (Schermerhorn, 2011). Therefore measuring the performance of an organisation remains at the very heart of human endeavour.

This chapter provides an introduction to the research area of interest, which is performance measurement (PM) in nonprofit organisations (NPOs). Therefore, the next section (section 1.2) provides the research background, which includes a description on the nature and diversity of the nonprofit sector (section 1.2.1) and a brief introduction to PM and PM frameworks (section 1.2.2). The next section (section 1.3) gives reasons for undertaking the research. This is followed by a statement of aims and objectives of the research (section 1.4), and for the sake of completeness, the corresponding research questions (section 1.5). The research questions have actually been derived from the synthesis of the literature (details in Chapter Three) and it is these research questions that legitimise the objectives of the study. The next section (section 1.6) covers the contribution of the research, both from a theoretical and practical standpoint. Section 1.7 states the limitations of the study (these limitations have been revisited in the final chapter). Finally, section 1.8 describes the structure of the thesis along with an explanation of how key elements of the thesis (thesis chapters and their subsections) are linked to one another in maintaining the flow of the thesis.

1.2. BACKGROUND TO THE RESEARCH

All organisations exist to achieve certain high level (strategic) objectives. In the case of nonprofit organisations (NPOs) most of these objectives, if not all, relate to serving the community to make some positive impact on them (Salamon & Anheier, 1997). No matter to which sector an organisation belongs (for-profit, public, nonprofit), it needs to measure how it is performing in achieving its strategic objectives. Formal PM Frameworks, such as the balanced scorecard (BSC) and the performance prism (PP) have been prescribed as models or tools that aid managers in developing appropriate performance measures to suit the strategic objectives of individual organisations.

Historically, performance reporting (hence PM) focused on the requirements of the owners (shareholders) of private sector organisations; the owners required financial information to ensure that they get the expected returns on the investments that they make. The managers who managed the organisations on behalf of the owners also needed financial information to take corrective action, when needed. Kaplan and Norton (1992) developed a more balanced perspective on performance reporting by including financial as well as nonfinancial dimensions (the BSC). The BSC is based on the premises that long-term financial sustainability of an organisation depends on how well it excels in certain other nonfinancial dimensions (Kaplan, 2008).

The main focus of formal PM frameworks is facilitating the generation of performance measures along predefined measurement areas (more technically, performance dimensions). For example, the BSC prescribes four performance dimensions while the PP prescribes five dimensions. While there is an underling theory for prescription of dimensions in each PM framework, there is no generalisable theory that underpins PM in the nonprofit sector (LeRoux & Wright, 2010). For example, the BSC for NPOs explains that having formulated logically connected strategic objectives (a strategy map) to achieve the mission, the managers should align each strategic objective to one of the four predefined performance dimensions (Learning and Growth, Internal Processes, Customer and Financial) and then choose at least one performance measure to relate to each strategic objective, given the four dimensions within which the managers have to work. One can argue that an organisation's strategy map explains how the strategies of an organisation causally relate to one another and how PM contributes in helping a NPO

to achieve its end goal (mission achievement). However, a strategy map can differ considerably from one organisation to another, thus restricting the generalisability of strategy maps even across similar organisations (e.g. organisations that provide the same type of service). Therefore the theory underling the BSC needs to be scrutinised and further developed, where necessary.

The operation and scope of NPOs can differ from country to country (e.g. USA versus New Zealand) as well as activity to activity (e.g. professional services versus community healthcare services). There are eight internationally accepted country clusters categorising NPOs and unsurprisingly, both New Zealand and Australia fall into the same cluster. This research focuses on generating and testing a theory that underpins PM in NPOs in New Zealand and Australia that provide *community healthcare* services.

1.2.1. The Nature, Diversity, and Growth of the Nonprofit Sector

In basic terms, a NPO is a tax exempt organisation. Therefore a NPO is a social entity that is organised for purposes other than generating profit, in the sense no part of the organisation's income is distributed to its members, directors, or officers (Barczak, Kahn, & Moss, 2006; Wu & Hung, 2008). Often NPOs are referred to as nonstock corporations to imply that a NPO does not have any owners holding equity (stocks) but can shadow a stock holding corporation in running day to day affairs. A NPO can however be formed as a corporation, an individual enterprise, an unincorporated association, a partnership, a foundation or a condominium (Brody, 2006). More recent legislation in certain jurisdictions (e.g. England and Wales) allow a NPO to function as a “charitable incorporated organisation” similar to a limited liability organisation but without being subjected to company law, when the aim of the organisation is to provide educational or charitable purposes (Morgan, 2010). Many NPOs are tax exempted and on occasions, they are also exempted from local taxes (sales or property taxes). Although NPOs can be formed as incorporated or unincorporated associations, the latter category of NPOs may not be eligible for tax *exempt status* in some jurisdictions.

Although the nonprofit sector is referred to by various names, each of these names emphasises only one aspect of NPOs, making light of what could be accepted as a

comprehensive definition (Salamon & Anheier, 1997). For example, the term *third sector* reminds us that NPOs differ from private and public sectors; the term *voluntary sector* emphasises the importance of the role being played by the volunteers in the sector; the term *independent sector* refers to the important role being played by these organisations as a *third force* that is independent from the government; the term *charitable sector* reminds us that NPOs look after the social wellbeing for no financial gain.

In addition to the above, there are some salient features of the nonprofit sector that distinguish it from the other sectors. *Culture* is the most prominent feature of NPOs; a warm, friendly and hospitable culture is pervasive in this sector. Since NPOs are mission driven organisations, the motive force of them is the *passion for the mission*. The mission of an organisation (nonprofit or otherwise) is a statement that clarifies the purpose of the organisation, indicating “*why it is doing what it does*” (Bryson & Alston, 2004, p. 67). All the members in a NPO—the executives, paid employees and volunteers—work dedicatedly to achieve the mission of the organisation. They gain satisfaction from accomplishing their tasks and seeing or believing that they contribute towards achieving the mission (Allison & Kaye, 2001; Barczak et al., 2006; Duque-Zuluaga & Schneider, 2008). In for-profit organisations the mission serves only as a means to an end (e.g. profitability) while in NPOs the mission serves as the end in itself (Barczak et al., 2006; Kanter & Summers, 1994).

There is a *scarcity of resources* in most NPOs, more so in small and medium sized NPOs; these organisations lack skilled decision makers (senior managerial personnel), other skilled personnel, and physical resources—to perform well (Allison & Kaye, 2001). Thus, these organisations are often resourced with less skilled and less experienced staff, which does not help in any way in measuring and managing performance.

In addition to the constraints mentioned above, NPOs face limitations in their flexibility to achieve their desired social impact (Kanter & Summers, 1994). This is mainly because NPOs are mission-driven and measuring the *success* or otherwise of achieving the mission (generally speaking, measuring the desired social impact) is very challenging (Bryson, 2011; Duque-Zuluaga & Schneider, 2008; Forbes, 1998; Kanter &

Summers, 1994). There are also political difficulties in designing measurement systems that accommodate the interests of various stakeholders in NPOs and this has also dampened the level of *mission achievement measurement* within the sector (Duque-Zuluaga & Schneider, 2008; Ritchie & Kolodinsky, 2003). Freeman (1984, p. 46) defines a stakeholder of an organisation as “*any group or individual who can affect or is affected by the achievement of the organization’s objectives*”. Studying organisational effectiveness in NPOs through standard PM tools also becomes complex due to their legal and financial status because of issues related to accountability; in spite of years of intense debate among scholars, there is very little agreement among them as “to whom” and “for what” NPOs are accountable (Andreaus & Costa, 2014). Yet another challenge in nonprofit PM is to generate quantitative measures on the intangible services that they provide in pursuing their unstructured goals (Barczak et al., 2006; Kanter & Summers, 1994).

For-profit sector’s success is mainly associated with profit and loss accounts, figures and the availability of myriads of quantitative measures to indicate profitability. This apparatus cannot create similar results for NPOs (Kanter & Summers, 1994; Moore, 2003). Unlike the for-profit organisations, NPOs have goals and services which are more intangible and harder to define (Kanter & Summers, 1994). Donors, beneficiaries (people receiving the services/community) and workers have different views about the quality of service of a NPO. In PM, this translates to measuring social value more than financial (economic) value. However, NPOs need financial stability to deliver their services to customers and for administrative tasks. Most NPOs depend heavily on external funding or philanthropic grants, such as donations from other organisations, governments or private donors. The NPO’s mission allows them to attract donors and government support because the donors believe that they donate for a good cause (Barczak et al., 2006). In the case of some grants, NPOs are subjected to the conditions that come with the grants.

Arguably, the two most important groups belonging to a NPO are the governing board and the volunteers (Renz & Anderson, 2014). The main responsibility of the governing board of a NPO is to make sure that the organisation serves the community requirements and the board contributes actively in achieving the success of the organisation (mission accomplishment). Unfortunately, as mentioned before, scarcity

of skilled people to govern NPOs impose a constraint on achieving organisational success (Allison & Kaye, 2001). Volunteers are a salient feature in NPOs and in fact, the board of directors themselves are volunteers as they are not paid as their counterparts in for-profit organisations (Salamon & Anheier, 1997); the voluntary employees contribute in providing the input to the NPO by way of knowledge and time in administrative or fieldwork for no financial return. In the absence of the volunteers, it would not be possible for a NPO to provide the services to the community.

The nonprofit sector is a fast growing sector (i.e. nonprofit sector growth rate > GDP growth rate) showing substantial economic presence in western countries throughout the world. Figure 1.1 depicts the growth rates of the sector in five western countries, based on the study reported by Salamon, Sokolowski, Haddock, and Tice (2013); Samalon et al. analysed the nonprofit sector in several countries using longitudinal secondary data covering the period 1990s through to mid-2000s. It is evident from Figure 1.1 that the GDP contribution of the NPOs in the five countries (unfortunately New Zealand was not included) outpaced the growth of the economy. Salamon et al. reported that the nonprofit sector contribution to the GDP in the five countries expanded at an average rate of 7.7% per year as against an average GDP growth rate of 5.9% (Salamon et al. reported that the value of volunteer time is not included the figures).

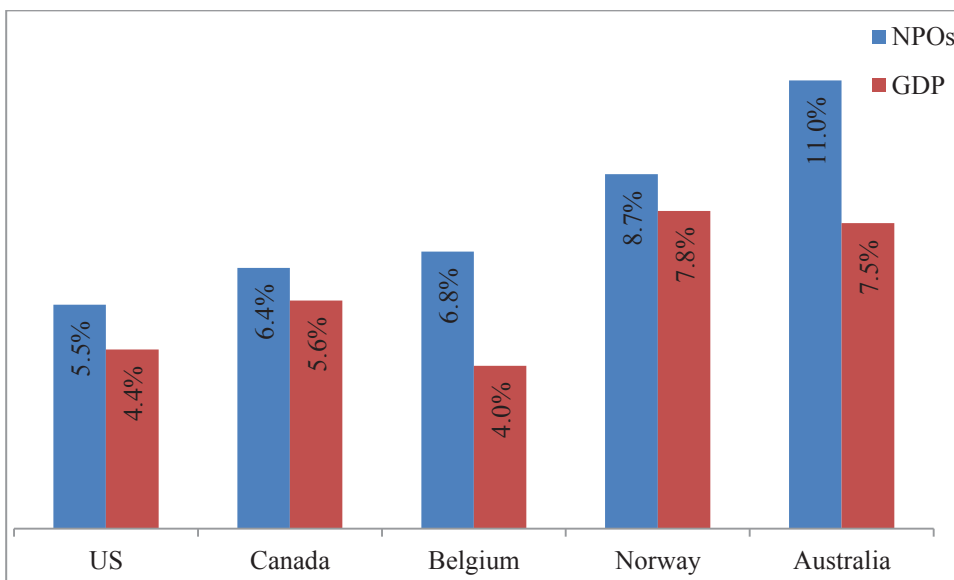


Figure 1.1: Nonprofit sector and GDP growth rates of five selected countries

With the growth of the sector, the question of accountability of the sector has become more and more pertinent and important. There is immense pressure from funders, governments, clients and the general public for NPOs to become more accountable for their actions (Andreas & Costa, 2014; Young, Bania, & Bailey, 1996). Therefore performance monitoring of these organisations has become a necessity rather than optional. Unfortunately, the sector needs a unified, reliable and valid PM framework to monitor the performance both at organisational level and sector/subsector level.

1.2.2. Performance Measurement and Performance Measurement Frameworks

A PM system can be defined as a “set of metrics used to quantify both the efficiency and effectiveness of actions” implemented to achieve the strategic goals of an organisation (Neely, Gregory, & Platts, 2005, p. 1229). The theoretical framework of a PM system facilitates the generation of performance metrics along predefined measurement dimensions. For each PM dimension, managers need to determine the appropriate set of measures to gauge the organisational performance (Kennerley & Neely, 2002). The measures are required to be derived from the strategy; they should be action-oriented and as a whole, reflect a balanced view of performance (Anthony & Govindarajan, 2007; Neely et al., 2005). A brief introduction to key existing PM frameworks follows.

The Balanced Scorecard

The BSC was developed by Kaplan and Norton (1992) from case research involving several leading US companies. The BSC is premised based on the “theory of the firm”, which posits that the operational objective of a firm is to maximise net benefits of future returns to the shareholders (owners). Kaplan and Norton found that if a firm is to sustain long-term financial performance, it should not just look at the financial dimension alone but should also look at customer performance, processes, as well as tangible and intangible inputs that feed the processes (Kaplan, 2008). Consequently, as mentioned earlier, the BSC entails four measurement perspectives (dimensions) aligned in a sequential order: Learning and Growth (process input), Processes, Customer, and Financial (Kaplan, 2008; Neely, 2007).

The BSC has come a long way from being a mere management accounting tool to a strategic decision making mechanism through a number of revisions and refinements (Anon, 2008; Kaplan, 2001, 2008; Lawrie & Cobbold, 2004). There is ample evidence in the literature to show that the BSC receives considerably greater practitioner and academia acceptance over alternative PM frameworks (details in Chapter Three).¹ Current versions of the BSC require performance measures to be aligned to strategic objectives in a causally logical manner representing the four PM perspectives of the scorecard (Hoque, 2014; Kaplan, 2008; Lawrie & Cobbold, 2004).

A nonprofit version of the BSC was published by Kaplan in 2001 (Kaplan, 2001). The primary difference between the theoretical underpinnings of the two frameworks is that whereas the regular BSC posits that it is the shareholder objectives that drive the strategy of the organisation, the nonprofit BSC posits that it is the organisation's mission that drives the strategy and not anything else (Kaplan, 2001; Niven, 2008). The researcher argues that although the mission driven notion is a sound argument, the underlying theory of the nonprofit BSC still requires refinements (details in Chapter Three). The following key shortcomings² of the nonprofit BSC are highlighted by the researcher: (a) lack of clarity on which customers (donors or clients) should represent the customer perspective; (b) lack of full consideration of key stakeholders (workforce satisfaction, which is vital to NPOs has been overlooked); (c) the role of financial outcomes; (d) inadequate explanation on causality; (e) difficulty in quantifying the overall organisational performance and the weights each measurement perspective (and its subparts) should carry; and, (f) lack of guidance on how a NPO could develop its own performance measures.

The Performance Prism

The PP was developed by Neely (2007) and his colleagues. The PP captures five facets on performance management: Stakeholder Satisfaction, Strategies, Processes, Capabilities, and Stakeholder Contribution. The salient facet of this model is Stakeholder Contribution which is not considered in many other PM systems (Forcada, Ramis-Pujol, & Cusumano, 2008). Although the PP pays more attention to the stakeholders, it suffers from several shortcomings: no clear link between the strategic

¹ For a quick look, see Figure 3.1 in Chapter Three.

² For a summary, see section 3.6.

indicators and the operational indicators, inadequate guidance as to how the prism should be practically implemented for performance improvement, weak theoretical grounding (the causal propositions are muddled), and lack of empirical support (Mendibil & MacBryde, 2005). Like the BSC, the PP does not prescribe how the overall performance could be quantified.

The Logical Framework

The logical framework (Kellogg Foundation, 2006) is a PM framework that has been designed to assess the social impact of any economic activity (profit driven or otherwise) that has a stake in the society. The framework consists of six performance dimensions arranged in the following causal order: Resources (inputs), Activities (processes), Outputs, Outcomes, and Impact. Although the logical framework looks beyond the inputs, processes, and outputs (the traditional systems notion), which is particularly relevant for NPOs, it is primarily being used to monitor performance at project level, rather than organisational level (Maas & Liket, 2011), thus limiting the possibility of being used for strategy implementation, monitoring and control.

Excellence Frameworks

Excellence frameworks such as the Baldrige Excellence Framework (USA) and the European Foundation for Quality Management (EFQM) Excellence Model are designed to produce multiple benefits to organisations: strategic PM, sharing and dissemination of best practices, and benchmarking (Grigg & Mann, 2008; Saunders, Mann, & Smith, 2007). Although these frameworks capture performance aspects related to key stakeholders (e.g. the workforce, customers/community, and donors) and promote leadership accountability, the researcher maintains that the universalistic nature of them (framework is claimed to be suitable for any organisation in any sector) is not in line with the researcher's fundamental argument that nonprofit management is different form for-profit management (details in section 2.3).

1.3. RESEARCH MOTIVATION

Despite a plethora of studies on PM in NPOs (section 3.4), the literature review revealed a number of knowledge gaps in PM as applied to NPOs (see section 3.6). More importantly, the researcher identified that these gaps could be addressed through a single research project (hence this thesis).

The BSC (by far most widely accepted PM framework by the academia and the practitioner community) provided a reasonable theoretical basis for nonprofit performance management but the literature synthesis resulted in unearthing several knowledge gaps (*these were briefly mentioned in the previous section*). The researcher also identified that alternative PM frameworks addressed some shortcomings of the BSC, but these had their own inherent weaknesses, which may partly explain why alternative PM frameworks have not been as widely accepted as the BSC.

The researcher also observed that the literature lacks large sample quantitative studies that examine the validity of the BSC in the nonprofit sector. In addition, there were very few studies that covered the Australasian nonprofit sector in healthcare (the healthcare sector is a reasonably homogeneous sector that favours case study type field research, which the researcher used in the theory development phase of her study).

All of the above meant that there was scope for a substantial theoretical contribution by way of theory development (developing a generalisable PM framework for NPOs using the BSC as the conceptual framework) and theory testing.

While the main motivation for undertaking the study was to address the knowledge gaps (not to mention the exciting fieldwork that invited mixed methods research) to make a significant theoretical and practical contribution, an added motivation was the phenomenal growth in the nonprofit sector (relative to the overall economy) in the Australasian region and the need for robust reliable PM systems to improve accountability of the sector. According to the Australian Bureau of Statistics (2015), from virtually a handful of entities 50 years ago, the Australian nonprofit sector has soared to 56,894 registered organisations at the end of June 2013 employing some 1,081,900 people, which is 9.3% of the Australian workforce (in addition to over 3.9

million volunteers) with an industry value addition of \$79.2 billion (also see Figure 1.1). Similar patterns exist in New Zealand (Sanders, O'Brien, Tennant, Sokolowski, & Salamon, 2008; Welch, 2007). The large number of potential respondents available also makes theory testing (validating the conceptual model) possible through appropriate statistical techniques.

1.4. AIMS AND OBJECTIVES OF THE STUDY

1.4.1. Research Aim

The aim of the research was to develop and test a generalisable PM system that could be used by NPOs to assess their strategic performance.

1.4.2. Research Objectives

The specific objectives of the study were as follows:

1. To determine how the BSC's performance dimensions (domains) should be augmented and operationalised to predict and explain the strategic performance of NPOs.
2. To develop a scientifically validated performance measurement instrument to test the validity of the BSC's performance dimensions (as augmented).
3. To test the theoretical causal relationships (i.e. the structural relationships) between BSC measurement domains and interpret the results from a theoretical and practical perspective.
4. To determine empirical weights for each performance dimension (and its subparts) to enable NPOs to assess their overall strategic performance.
5. To compare and contrast the researcher's PM framework with the existing PM frameworks—in particular the BSC, which enabled the researcher to initiate her theory development.

The general objective of the study is to develop guidelines for practitioners in NPOs to help them to assess the strategic performance of their organisations.

1.5. RESEARCH QUESTIONS

Formulation of research questions requires a careful synthesis of the literature (Chapter Three) to identify the knowledge gaps. This chapter (mainly section 1.3) provided only a concise summary of the knowledge gaps. The researcher has provided a full justification of her research questions in Chapter Three (in particular, see section 3.6). For the sake of completeness, the research questions are summarised in this chapter as follows.

RQ1. What set of PM dimensions constitute an integrated PM system for NPOs?

RQ2. What are the operational definitions of these PM dimensions?³

RQ1 involves formulating a theory (using the nonprofit BSC as the initial conceptualisation) that predicts and explains how the Mission of a NPO drives its strategies to achieve Stakeholder Outcomes to achieve its Mission. The constructs of the researcher's theoretical framework (model) represent the PM dimensions related to nonprofit PM. The second research question relates to operationalisation of the PM dimensions of the researcher's theoretical framework. After answering RQ2, the next step would be to test the hypothesised theoretical relationships (RQ1) between the constructs. This leads to third research question.

RQ3. How do the PM dimensions relate theoretically to one another in explaining the achievement of strategic objectives of a NPO?

After scientifically validating the researcher's PM framework (i.e. answering the RQ3), it is important to guide the practitioners/managers how to use the PM model to assess

³ The term *operational definition* refers to making abstract concepts (constructs) concrete by defining how these concepts are measured (in concrete terms) within the causal structure within which the constructs are related. Testing and validating a survey instrument that captures the abstract constructs fulfils this requirement (Cooper & Schindler, 2014).

the overall performance (in numerical terms, as a score) of their organisation. This leads to fourth research question.

RQ4. What weighting should managers give to each PM dimension and its subparts in assessing the overall strategic performance of a NPO?

The overall research question of the study is:

What are the performance dimensions that underpin PM in a NPO and how are these dimensions causally related?

The reader should note that the researcher uses the terms “performance dimensions” and “PM dimensions” interchangeably throughout this thesis. The nexus between the research questions and research objectives are shown in Figure 1.2.

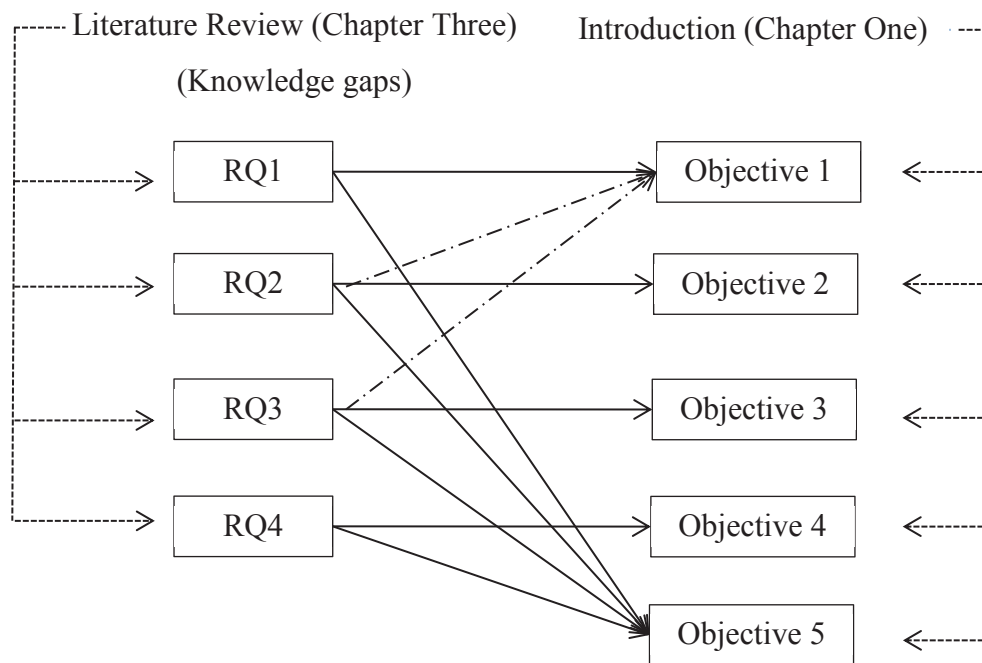


Figure 1.2: The nexus between research questions and research objectives

1.6. RESEARCH CONTRIBUTION

Theoretical Contribution

The major theoretical contribution of this study is adding new knowledge by way of deriving a scientifically validated PM framework for NPOs. The study addresses several knowledge gaps in the nonprofit BSC and therefore it helps academia expand current knowledge on the theoretical basis of the BSC dimensions. This is because the core proposition of the nonprofit BSC “*mission drives the strategy to align tangible and intangible resources (organisational leaning and growth) to feed processes to achieve mission-related outcomes*” is retained in the researcher’s theoretical framework (hence the PM framework that she developed).

The study is unique because it is the first of its kind that attempts to examine the generalisability of any PM framework across the nonprofit sector. The reader should note that the study was limited to the Australasian healthcare sector due to theoretical and practical reasons. The researcher uses fieldwork to: (a) improve the theoretical framework of the BSC (via small sample qualitative data), (b) test the operationalisations of the performance dimensions of her PM framework as well as the hypothesised theoretical relationships between the performance dimensions, and (c) to determine empirical weights for each performance dimension and its subparts ((b) and (c) via large sample quantitative data). The study also uncovers contextual information of the nonprofit healthcare sector in the Australasian region (e.g. issues underlying PM, key resources required, key stakeholders) which may also lead to further research.

Practical Contribution

The major practical contributions of this study are: (a) the researcher’s PM framework, (b) the PM instrument for NPOs, (c) the scoring system developed by the researcher, and (d) general guidelines given by the researcher on how to use her PM framework to improve the strategic performance of a NPO. Each of these is explained in turn.

The researcher’s PM framework, which has been tested empirically, explains how mission-related outcomes (stakeholder satisfaction) are achieved through the organisation’s strategy by aligning its tangible and intangible resources to feed its processes. Thus managers can use the researcher’s PM framework to explain successful

strategy implementation and unsuccessful strategy implementation (academia may view this as a virtue of a good theory). The PM instrument developed by the researcher to test her PM framework can be used by practitioners to self-examine how their organisation is performing in the leading indicators of mission-related outcomes as well as the mission-related outcomes themselves. Parent organisations may use the PM instrument and the associated scoring system (an explanation follows) to benchmark its umbrella organisations to help them learn from each other. The scoring system developed by the researcher to weight each performance dimension (domain) and its subparts to determine the overall organisational performance should further help practitioners in self-examination of their organisation's performance. The guidelines given by the researcher on how to use her PM framework to improve strategic performance includes an explanation of what her PM framework explains, how to self-examine strategic performance, and how they could identify performance metrics on each of the performance dimensions entailed in the PM framework. The benefits of these guidelines to the practitioner could be substantial.

1.7. LIMITATIONS OF THE STUDY

The study uses data collected from the Australasian nonprofit healthcare sector only. This applies to both theory refinement, operationalisations of PM dimensions (based on small sample qualitative data), and theory testing (based on large sample quantitative data). Generalisations of this study should be made with caution, even within nonprofit healthcare. The nonprofit sector is a very diverse/heterogeneous sector⁴, although confining the study to the healthcare subsector reduces the heterogeneity substantially.

Due to resource constraints (time, funding etc.) the researchers' PM framework and the PM instrument could not be tested for practical usability/validity in a longitudinal type of a study. For this reason this research requires a follow-up study to reap its full practical benefits.

In addition to the above, the study does not examine in any great detail to what extent the findings can be used in public sector performance management (public sector performance management is not covered in this study). Both the public sector and the

⁴ For a detailed examination of the sector see Salamon and Anheier (1997).

nonprofit sector provide social goods (sometimes goods provided by one sector happens to be close substitutes of the goods provided by the other sector) and although the PM fraternity treat performance management in both sectors as being similar—for example Kaplan (2001), the inventor of the nonprofit BSC asserts that the scorecard can be used by public enterprises also—certain differences between the two sectors may mean that the findings of this study could be less useful to the public sector.

1.8. THE STRUCTURE OF THE THESIS

The rest of the thesis is structured as follows.

Chapter Two provides an overview of NPOs. This begins with a historical background of NPOs (how and why they evolved) and a comprehensive definition and classification of NPOs. This is followed review of corporate governance and accountability literature, particularly in relation to NPOs. The researcher then introduces the nonprofit sector in Australasia along with an explanation of how healthcare NPOs are classified in Australia and New Zealand. The researcher then elaborates on the challenges of nonprofit PM introduced in Chapter One. In essence, Chapter Two provides the contextual information on NPOs required for the research.

Chapter Three provides a synthesis of the primary literature relevant to achieve the aims and objectives of the study. The key areas covered in this chapter includes review of the literature on the theoretical underpinnings of the BSC and other relevant PM frameworks, studies that examined various facets of the validity of the BSC (BSC implementation studies, conceptual papers, and empirical studies), identification of knowledge gaps, and the research questions.

Chapter Four covers the research methodology. This chapter begins with an overview of paradigms available to conduct social research. This is followed by coverage of relevant literature on mixed methods research and justification of the paradigm chosen by the researcher. The chapter then presents a conceptual model (based on the extant literature) to commence fieldwork. This is followed by a description as to how the conceptual model was developed into a testable theoretical model of PM by conducting fieldwork (small sample qualitative data collection in a case study setting). This

description includes how the performance dimensions were operationalised and how the hypotheses (theoretical relationships between performance dimensions) were formulated (these relate to RQ1, RQ2 and RQ3). Next, the chapter describes the development of the quantitative survey questionnaire that was used to test the operational definitions of the performance dimensions and theoretical relationships (hypotheses). This is followed by a description of how the survey was administered and data were analysed to examine the hypothesised relationships (the data analysis provided concrete answers to RQ1 through RQ3).

Having dealt with the methodology associated with the theoretical model (known as the main model in this thesis), the researcher developed an ancillary theoretical model to fit survey data to the model, to estimate empirical weights for each performance dimension and its associated subparts (this relates to RQ4). Finally, the researcher looks at her methodology in retrospect to compare it against the quality criteria prescribed in mixed methods research. Parts of this chapter have been published in three conference proceedings. An overview of the study was presented at the 11th ANZAM Operations, Supply Chain and Services Management Symposium (Soysa, Jayamaha, & Grigg, 2013b), description of the adopted methodology was presented at the 2013 Performance Management Association of Australasia (PMAA) Conference (Soysa, Jayamaha, & Grigg, 2013a) and the development of the conceptual model was presented at 12th ANZAM Operations, Supply Chain and Services Management Symposium (Soysa, Jayamaha, & Grigg, 2014).

Chapter Five reviews the literature on tools and techniques used in *positivistic* social and behavioural science research to test theoretical models containing abstract concepts; performance dimensions are abstract because they are not directly observable. The purpose of this chapter is to better inform both the researcher and the reader of the specific psychometric (measurement of abstract concepts) principles, statistical theory and modelling relevant to the study.

Chapter Six covers results and discussion of the case study. Upon showing the relevant details of the case study organisations ($n = 9$) and their respondents, the chapter shows the qualitative data collected from the respondents. These qualitative data are first presented as ‘fragmented responses’ under each theme (performance dimension) and

subtheme. The chapter then engages in a detailed discussion on each theme and subtheme, using quotes provided by the respondents. This is to shed more light on what each theme represents, from a PM perspective. Thereafter, the chapter provides the operational descriptions of each performance domain (along with the final theoretical model), which enabled the researcher to translate these descriptions in an alternative format to develop the quantitative survey questionnaire. The chapter then discusses the theoretical and practical implications of the case study findings. In effect, the chapter answered RQ1 and RQ2, allowing the researcher to achieve the first two research objectives. A significant portion of the contents in Chapter Six will appear in a forthcoming article in the *TQM Journal* (Soysa, Jayamaha, & Grigg, 2016b).

Chapter Seven discusses the descriptive statistics of the quantitative data collected. The statistics reported in this chapter facilitate the discussions provided in the next chapter. Chapter Eight covers the results and discussion on the hypothesised theoretical relationships in the PM framework as well as the parameter estimates of the ancillary theoretical model. Operational definitions that were not supported by data are also highlighted. Although results and discussion provided in Chapter Six addressed the first two research questions (RQ1 and RQ2), it is the results and discussion in Chapter Eight that comprehensively answer these research questions as well as the third research question (RQ3); in answering the first three research questions, the researcher achieved the third research objective. A substantial portion of the aforementioned components of Chapter Eight is in the final stages of review (minor revisions) in the *Total Quality Management & Business Excellence* journal.

The parameter estimates of the ancillary theoretical model along with the empirical weights of each performance dimension and its subparts are also presented and discussed in Chapter Eight. This answered the fourth research question (RQ4) to enable the researcher to achieve the fourth research objective. The implications of the findings on the ancillary model to management practice have also been discussed in the chapter. The development of the ancillary model as well as the results and discussion were presented at the 2016 Performance Management Association (PMA) Conference (26-29 June) held in Edinburgh, Scotland, UK (Soysa, Jayamaha, & Grigg, 2016a). Towards the end of Chapter Eight, the researcher compares her PM framework with the BSC and

Excellence Models. This achieved the fifth research objective. See Appendices F and G for all publications related to this study.

Chapter Nine, the final chapter, covers the conclusions of the study in relation to the research questions and research objectives. A justification of theoretical and practical contributions of the study has been provided, along with potential future research directions.

The links between the research questions, research objectives, key findings and the thesis chapters are depicted in Figure 1.3.

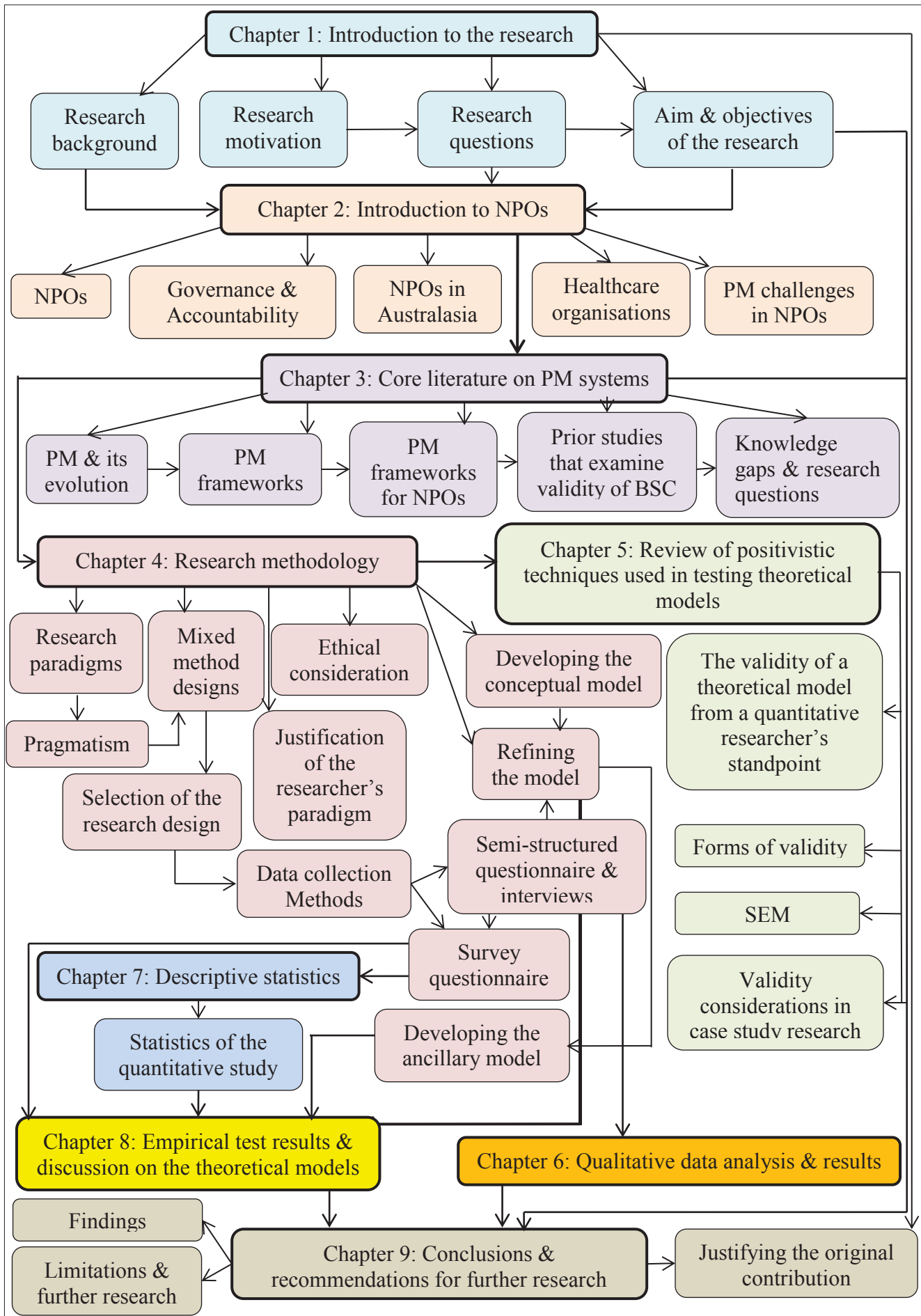


Figure 1.3 : The structure of the thesis

CHAPTER TWO

INTRODUCTION TO NONPROFIT ORGANISATIONS

Write what you know. That should leave you with a lot of free time.

Howard Nemerov

2.1. INTRODUCTION

This chapter provides the necessary theoretical grounding on nonprofit organisations (NPOs) for the researcher's study. In section 2.2 the researcher provides an overview of the origin of the modern NPO. This section also provides an operational definition of NPOs to reliably distinguish NPOs from for-profit and government organisations. Widely accepted theories that explain the existence of NPOs alongside for-profit and government organisations are also explained in section 2.2. The next section (section 2.3) covers the complementary concepts *corporate governance* and *accountability*, particularly in relation to NPOs. The relevant competing theoretical frameworks on corporate governance—agency theory, stakeholder theory and resource dependence theory—are also covered in section 2.3, again particularly in relation to NPOs. Section 2.4 reviews the nonprofit sector in Australasia while section 2.5 describes how the nonprofit healthcare sector is structured in the region. Section 2.6 emphasises the challenges NPOs face in measuring and monitoring strategic performance. Finally, section 2.7 concludes the chapter by highlighting the key points covered in the chapter.

2.2. NONPROFIT ORGANISATIONS: THE GENESIS, DEFINITION AND ECONOMIC THEORIES

2.2.1. The Genesis of NPOs

In 1601 the British government passed the ‘Charitable Uses Act 1601’, which stipulated the purposes for which private cash could be used for the supply of public goods, with *social recognition* tagged into it. Until that time, it was rare to direct money to any other purpose but the benefit one's heirs, upon one's death (Cordery & Baskerville, 2007). The British Charter of Rights (1689) passed by the British parliament stipulated the rights and liabilities of the British subjects to petition the monarch without fear of any repercussion. A similar bill was passed in the United States, namely the US Bill of Rights (1791), to limit the power of the US federal government to enable the US citizens the freedom of enjoying the natural rights, as they are known today. These include, the freedom of forming associations, the freedom of speech, the freedom of the press to publish news and many other rights, which we take for granted today. These statutes are regarded as the precursors of nonprofit associations. The notion of the ‘*nonprofit organisation*’ as a unified *sector*, goes back to 1970, although some educational and religious organisations such as the churches have a long history and tradition of serving the public for nonfinancial gain (Hall, 2005). Religion has always been a facilitator for nonprofit activities because ‘giving’ is a deep rooted concept in the Christian, and many other faiths. The religious community has an impact on education, health care, human services, arts and advocacy. Religious organisations have created a network of ‘faith-based schools’, hospitals and human services groups (Hall, 2005).

When it became legal to do so, people started forming associations, foundations and institutions to deliver human services, promote grass-roots economic development, prevent environmental degradation, protect civil rights and pursue a variety of other objectives formerly unattended or left out by the governments (Salamon, 1994). Salamon observed that the growth of the nonprofit sector springs from the pressures of individual citizens, outside institutions and the governments themselves. The emergence of NPOs brought about not only a distinct set of social and technological changes but also confidence in the capability of the government. Broad historical changes have thus opened the way for alternative institutions that can respond more effectively to certain human needs—in economic terms, provision of ‘collective goods’

(Salamon, 1994; Weisbrod & Dominguez, 1986). With their small scale, flexibility and capacity to engage grass-roots energies, NPOs become ideally suited to fill the resulting gap (Salamon, 1994).

2.2.2. Defining NPOs

As mentioned in the Chapter One, the differences between NPOs and public and private organisations give rise to a need for a structural-operational definition of NPOs. The characterisation and definition of the nonprofit sector as well as classification of organisations belonging to the sector (this eventually resulted in an international classification of NPOs) are a result of pioneering work done by Lester Salamon and his associates at the Johns Hopkins University (JHU), USA. In one of the key research projects, Salamon and his associates examined the structure and characteristics of NPOs in over 13 countries to understand the structure of NPOs in the US and elsewhere (details in Salamon & Anheier, 1997). This project published country-specific findings through a series of working papers; for the New Zealand country study, see Tennant, Castle, O'Brien, Salamon, and Sanders (2006); for the Australian country study, see Lyons, Hocking, Hems, and Salamon (1999). The JHU study identified five key structural and operational features of a NPO (Salamon & Anheier, 1997). Each of these is described in turn.

The first key feature is that NPOs are *organised*. The term 'organised' was used by JHU researchers to mean that NPOs are highly structured in that they are methodical and efficient in their arrangement and operations, whether or not they are legally registered. The second key feature is that NPOs are *private*. The term 'private' was used to mean that the nonprofit sector is not 'another arm of the government', although NPOs are meant to provide basically the same collective goods that a government usually provides, for which, most NPOs receive government funding (Salamon & Anheier, 1997). The third key feature is that NPOs are *not profit distributing*. According to Salamon and Anheier, this feature is the most pervasive characteristic of a NPO; by law, a NPO should exist for purposes other than generating profit; no part of the organisation's income is allowed to be distributed to its members, directors, or the officers. The fourth key feature is that NPOs are *self-governing*. This refers to the fact that NPOs have their own ways and means of pursuing their operations and achieving

targets. The fifth and final feature is that NPOs are *voluntary*. NPOs emphasise its membership and the community that donation of money or time are not required by law but desired for achieving a specific mission (Salamon & Anheier, 1997). The aforementioned final feature of a NPO highlights the importance of the role being played by the volunteers in the nonprofit sector.

The aforesaid structural-operational definition has been tested for descriptive validity in every country covered by the project (41 countries including Australia and New Zealand). The objective of these studies was to determine whether or not the five factor definition of Salamon and Anheier (1997) is wide enough a definition to embrace the different range of organisations that are considered to be part of the nonprofit sector worldwide; in particular, the researchers verified that the definition distinguishes NPOs from for-profit and government organisations at a global level. The tests found that although there are certain country specific characteristics in the sector, the said structural-operational definition of NPOs is robust enough to be used as an international standard to classify NPOs.

2.2.3. The International Classification of NPOs

According to the *International Classification of Nonprofit Organisations* (ICNPO), NPOs are classified under 12 main activity groups. The ICNPO subdivides the 12 main activity groups into 30 subgroups (Table 2.1). Even though subgroups can be further divided based on activity, stratification at activity level is not incorporated in the ICPNO, because NPOs differ from one jurisdiction to another at activity level (Salamon & Anheier, 1997).

Table 2.1: International Classification of Nonprofit Organizations (Salamon & Anheier, 1997)

Groups (Code)	Subgroups (Code)
Culture and recreation (1)	Culture and arts (1 100)
	Recreation (1 200)
	Service clubs (1 300)
Education and research (2)	Primary and secondary education (2 100)
	Higher education (2 200)

Groups (Code)	Subgroups (Code)
	Other education (2 300)
	Research (2 400)
Health (3)	Hospitals and rehabilitation (3 100)
	Nursing homes (3 200)
	Mental health and crisis intervention (3 300)
	Other health services (3 400)
Social services (4)	Social services (4 100)
	Emergency and relief (4 200)
	Income support and maintenance (4 300)
Environment (5)	Environment (5 100)
	Animals (5 200)
Development and housing (6)	Economic, social and community development (6 100)
	Housing (6 200)
	Employment and training (6 300)
Law, advocacy and politics (7)	Civic and advocacy organizations (7 100)
	Law and legal services (7 200)
	Political organizations (7 300)
Philanthropic intermediaries and voluntarism promotion (8)	Philanthropic intermediaries (8 100)
International Activities (9)	International activities (9 100)
Religion (10)	Religious congregations and associations (10 100)
Business and professional associations, unions (11)	Business and professional associations, unions (11 100)
Not elsewhere classified (12)	Not elsewhere classified (12 100)

Through a study involving 41 countries, the researchers at the JHU identified eight country clusters that characterise NPOs: Welfare Partnership, Anglo-Saxon, Nordic, Eastern European, Asian Industrialised, Latin American, African, and Other Developing (Salamon & Sokolowski, 2004). In particular, the JHU researchers considered the basic structure, the size of the sector, level of volunteer participation, revenue, level of development, social and political history, geographical region, culture and traditions of each country, in creating the eight country clusters. Needless to say, New Zealand and Australia fall into the *Anglo-Saxon* cluster.

2.2.4. Theories That Explain the Existence of NPOs

There are three well-known economic models in the nonprofit literature (e.g. Bowman, 2011; Steinberg, 2006) which collectively explain why NPOs exist. In his pioneering work, Weisbrod (1975) argued that the nonprofit sector exists because of the failure of the government sector to supply a certain quantity of “higher levels of services” that the market (the private sector) fails to provide. Weisbrod argued that the government has capacity limitations and therefore will not undertake to supply the demand for all of the goods and services that the private sector (the market) fails to supply. Thus the *market failure model* and the *government failure model* were the earliest models that were developed to explain why a nonprofit sector exists (Bowman, 2011; Steinberg, 2006). Subsequently, Salamon (1987) added a third dimension, the failure of the volunteer sector, which gave rise to the volunteer failure model (Bowman, 2011; Steinberg, 2006). Each of the three models is explained in turn.

The *market failure model* argues that the market fails to supply certain goods and services in three situations—under-provision, over-exclusion, and contract failure—which causes the government sector to respond to the said failures (Steinberg, 2006). Under-provision is said to occur when the for-profit sector fails to produce a certain good or service needed by the public due to lack of commercial viability. Over-exclusion is said to occur when the market excludes borderline non-purchasers of public goods by overpricing the good/service. Finally, the contract failure is said to occur when the consumers fail to accurately determine the quantity and/or quality of the product/service produced for them (Hansmann, 1980; Steinberg, 2006).

The *government failure model* explains situations in which the government fails to adequately respond to the three types of market failures described above. A discussion on these situations is beyond the scope of this study (for details see Steinberg, 2006, pp 122-125). The *voluntary failure model* advanced by Salamon (1987, pp 39-42) explicates four situations—*philanthropic insufficiency*, *particularism*, *philanthropic paternalism*, and *philanthropic amateurism*—under which the nonprofit sector fails to provide a public good, which provides an opportunity for the for-profit or the government sector to provide that good.

2.3. GOVERNANCE AND ACCOUNTABILITY IN NPOS

Governance and accountability are two terms that go hand in hand, whichever way one looks at an organisation, because organisations need a governance structure to hold them accountable for their action. Being nonprofit makes an added dimension to governance and accountability (Hoque & Parker, 2015; Renz, 2007).

2.3.1. Nonprofit Governance

The researcher distinguishes nonprofit governance from corporate governance. The latter relates to governance of corporations (firms) that have ‘share ownership’, with a clear separation of ownership and management (Solomon, 2010). While the literature on nonprofit governance has been cross-fertilised with the corporate governance literature, especially from the “agency theory” on one extreme and the “stakeholder theory” on the other extreme (Hoque & Parker, 2015; Renz, 2007; Renz & Anderson, 2014), for the purpose of the researcher’s study, conceptual debates on the merits and demerits of different corporate governance frameworks have been set aside.

In this thesis, the researcher uses the definition of governance given by Renz and Anderson (2014, p. 18): “governance is the systems and procedures concerned with ensuring the overall direction, control, and accountability of the organisation”. According to Speckbacher (2008), governance covers all the important strategic activities such as goal setting and direction, making policy and strategy decisions, managing and monitoring organisational performance, and ensuring overall accountability. Governance is a term that is invariably associated with the governing board of an organisation (in NPOs as well as other organisations). A governing board is ultimately accountable for all the actions that come under the organisation, whether or not those actions are based on the board’s decisions. The members of the governing board have certain legally enforceable duties and obligations, by virtue of their membership on the board (Renz, 2007). However, it is not uncommon for staff members in larger organisations to be a part of the governance process.

The chief executive officer (CEO) or the chief operational officer of the organisation plays a very active role in the governance of an organisation (Renz, 2007). The CEO of a NPO provides strategic leadership and direction to the organisation by working with

the board, in ensuring that the strategies are formulated and implemented to achieve the vision and the mission. Renz asserts that it is important to renew the link between the services provided by the NPO and the requirements and interests of its target community on a continuous basis, for good governance and organisational success. It is the responsibility of the governing board to ensure that the services provided by the organisation are valued by the clients and the constituents in ways that are consistent with the organisation's core values and mission. While nonprofit governance involves making high-level decisions to cause efficient and effective services, contemporary NPOs are under pressure from external stakeholders to be more accountable and transparent in the organisation's activities (Ostrower, 2007). This reshapes how contemporary NPOs are managed and governed.

The researcher covers three key competing theories on nonprofit governance relevant to her study (especially the first two given below), although these theories have evolved in fields outside nonprofit management such as 'business ethics'.

2.3.1.1. Agency Theory

Agency theory highlights the so-called "agency problem" which arises due to separation of owners (principals) from managers (agents) of the company (Renz & Anderson, 2014; Solomon, 2010). The agency theory argues that 'egotism' on the part of the agents, cause them to act in their own self-interests at the expense of maximising shareholder value (Renz & Anderson, 2014; Solomon, 2010). Brennan and Solomon (2008) show that researchers on corporate governance who subscribe to agency theory focus on resolving conflicts of interest between the principal and agent. Renz and Anderson (2014) observe that many researchers have attempted to use agency theory to explain nonprofit governance, but with limited success. In particular, Renz and Anderson observe that the researchers had difficulty in distinguishing between the principals (owners) and the agents of a NPO in the first place. Renz and Anderson observe that some corporate governance researchers who used the agency theory assumed that the board of directors are the principals and that the executives are the agents while other researchers assumed that the board of directors are also agents because the board governs the organisation for the donors, clients and the wider community.

Agency theory is important to performance measurement because, traditional management accounting (hence performance measurement systems such as the balanced scorecard that have been developed by researchers belonging to the management accounting discipline) and agency theory share the common assumption that the role of a firm is to maximise shareholder value (Freeman, Harrison, Wicks, Parmar, & Colle, 2010). Hoque and Parker (2015) highlight that the nonprofit distributing nature associated with NPOs question the currency of agency theory in a nonprofit context, although it is possible that self-serving members of a NPO can misdirect funding (donations) for self-gain, to bring some credence to agency theory.

2.3.1.2. Stakeholder Theory

Stakeholder theory was proposed forty years ago as a counter theory to better explain certain key problems associated with the agency theory and the profit maximising prophesy that were dominant at the time (Freeman, Harrison, Wicks, Parmar, & De Colle, 2010; Freeman, 1994). According to Freeman et al. (2010), who were instrumental in advancing stakeholder theory, stakeholder theory provides an alternative—and supposedly richer—explanation to three key questions, which *approximately translate* to the following inquiries: how does a business create value and promote trade, in a dynamic environment?; what does business ethics mean in modern capitalism?; given the environments in which contemporary business operate, what is management?

Although reviewing the evolution of stakeholder theory is beyond the scope of this study (for a synthesis on the key phases of the evolution of stakeholder theory see Elias, Cavana, & Jackson, 2002), the researcher highlights four formats of the stakeholder theory explicated by Donaldson and Preston (1995), due its relevance to the balanced scorecard, a key theoretical framework that underpins the researcher's study. The '*descriptive stakeholder theory*' views a corporation as a "constellation of cooperative and competitive interests possessing intrinsic value" (p. 66); '*instrumental stakeholder theory*' highlights the importance of stakeholder management to achieve performance goals, on the premise that corporations that are better at stakeholder management achieve better long-term performance outcomes in traditional performance indicators (e.g. profitability, growth) than the ones that look at the principals' interests alone;

'normative stakeholder theory' argues that stakeholders have “legitimate interests” in “corporate activity” and that the interests of all groups of stakeholders carry “intrinsic value”; finally, *'managerial stakeholder theory'* argues that corporations require “simultaneous attention to the legitimate interests” of all key stakeholders (p. 67). Renz and Anderson (2014) observed that stakeholder theory has been used in nonprofit governance studies for mapping internal and external stakeholders and to categorise and explain the roles being played by the boards (Renz and Anderson show the boards operate at the interface between the internal environment and the external environment of a NPO).

2.3.1.3. Resource Dependence Theory

Resource dependence theory (Salancik & Pfeffer, 1978) argues that organisations will engage with fruitful relationships with the actors in the external environment so as to reduce dependency on resources to mitigate risk and uncertainty. A view of the literature on nonprofit governance (e.g. Brown, 2005; Heimovics, Herman, & Coughlin, 1993; Malatesta & Smith, 2014) indicated that resource dependence theory has been accepted by most researchers for its relevance in explaining nonprofit governance, especially in explaining the roles being played by boards in reducing risk and uncertainty.

2.3.2. Accountability

Nonprofit or otherwise, accountability is a term that goes hand in hand with (good) governance. This is because in *management*, accountability is a natural consequence of delegating authority to a person or a body to execute a certain task; in the case of a for-profit organisation, the responsibility of the board is to manage the organisation on behalf of its principals (Albareda, 2008). In this section, the researcher uses the term accountability to mean the accountability of the board of directors (often known as ‘board accountability’) in a NPO, in relation to the corporate governance literature.

Scholars have proposed several definitions of accountability. Cornwall, Lucas, and Pasteur (2000) stressed that accountability has both an external dimension as well as an internal dimension, because accountability is about “being held responsible” (the external dimension) by others and about “taking responsibility” for one’s own conduct

(the internal dimension). Gray, Owen, and Dams (1996) define accountability as “the duty to provide an account (by no means necessarily financial) or reckoning of those actions for which one is held responsible”. Accountability not only implicates a promise to perform, but also a legal responsibility to provide an account for it (Cornwall et al., 2000). Costa, Ramus, and Andreus (2011) asserted that accountability contains two interconnected but dissimilar responsibilities: (i) the obligation to undertake certain activities and (ii) the obligation to be responsible for those activities. Although it is clear to whom the board of a for-profit organisation is accountable, the same cannot be said about NPOs, because traditional notions of governance are based on the principal-agent relationship, which is not directly applicable to a NPO (Hoque & Parker, 2015; Viader & Espina, 2014; Renz & Anderson, 2014). A discussion on the complexity of nonprofit board accountability follows.

Nonprofit governance research identified two accountability domains of NPOs: the upward accountability and downward accountability (Ebrahim, 2003; Ebrahim, Battilana, & Mair, 2014; Flack & Ryan, 2005; Ospina, Diaz, & O’Sullivan, 2002). *Upward accountability* of a NPO refers to the board’s accountability to the trustees, donors and the government while *downward accountability* refers to accountability to the beneficiaries, communities, and partners. It is well documented in the nonprofit governance literature (e.g. Ebrahim, 2003; Flack & Ryan, 2005; Laratta, 2010) that downward accountability is nearly always subordinated in favour of upward accountability (particularly, the board’s accountability to the donors and the government). Another dichotomy associated with board accountability in the accounting literature is *administrative and procedural accountability* versus *programmatic accountability* (Najam, 2000). The process leading the contracting relationship—internal audits requirements, billing requirements and obligations with government laws—is referred to as administrative and procedural accountability while the capacity of the service delivered by the organisation under government contract is referred to as programmatic accountability (Najam, 2000). According to Edwards and Hulme (1996), programmatic accountability is more complex than administrative and procedural accountability because the former often translates to ‘accounting’ rather than accountability, leading NPOs to count services and clients rather than focusing on their performance. NPOs respond to the concerns of trustees, donors and the government through procedures, policy documents and funding requirements (Carman, 2009).

These activities not only document program activities and efforts but also become an integral part of performance-based accountability, because they provide evidence that a program is being delivered as planned (Love, 2004, p. 96). In particular, managers have to convince the donors and the board that funding has been effectively utilised, the clients are being served, and that programmes are being implemented as planned.

On the other hand, NPOs are not officially accountable to beneficiaries (Chisolm, 1995). Kilby (2006) asserts that although the literature on nonprofit governance highlights the importance of the relationship between NPOs (the boards) and their beneficiaries, there is no clear accountability path to the beneficiaries. More importantly, the literature on NPO accountability recognises several techniques that capture upward accountability but specifies few on downward accountability (Einfeld, 2001; Kearns, 1996). According to Fry (1995), downward accountability is more complex than upward accountability since it has two dimensions: a feeling of responsibility (internal dimension) and responsible behaviour (external dimension).

According to scholars, governmental regulation and oversight of NPOs have grown considerably since the 1990s, in an effort to ensure their accountability to the public (Austin, 2003; Najam, 2000; Tremblay-Boire & Prakash, 2015). This has created a division among scholars—some claiming the pressure for accountability as a disadvantage for NPOs (Vinson & Morley, 2000), while others consider it as a benefit (Laratta, 2010). Many scholars claim that because of accountability pressures, the focus of NPOs has deviated from achieving their mission to maintaining financial stability and conforming to regulatory procedures (Austin, 2003; Lawry, 1995; Najam, 2000).

2.4. NONPROFIT SECTOR IN AUSTRALASIA

While the ICNPO is a useful scheme to classify NPOs, Salamon and Anheier (1997) opined that the inherent attributes and the function of the nonprofit sector in any country is eventually shaped by the entire pattern of social, economic and political development of that country. Anglo-Saxon countries (section 2.2.3) including New Zealand and Australia share a common historical connection with the Anglo-Saxon political and legal tradition (Salamon & Sokolowski, 2004). The Salient features of the nonprofit sector in Anglo-Saxon countries are: a sizable nonprofit sector operation with very low

contribution from the government (only about a third of total revenue, on average, per country), greater voluntary participation (twice as much relative to other countries), and greater philanthropic support in the form of cash and revenue (Sanders et al., 2008). The other important feature of the nonprofit sector in Anglo-Saxon countries is the greater emphasis on community service functions (e.g. healthcare) relative to other service functions.

New Zealand and Australian NPOs represent a significant economic presence in the respective countries. These organisations are important to the two countries, both socially and politically (Scaife, McDonald, Williamson, & Mossel, 2015; Tennant et al., 2006). According to the United Nations' handbook on the nonprofit sector (Salamon et al., 2013), the contribution of the nonprofit sector to the gross domestic product (volunteer workforce contribution included) in New Zealand and Australia are 5.3% and 4.9% respectively, and more importantly, growing. NPOs engage a workforce that exceeds the combine workforce of the transportation and finance sectors—two sectors that represent major industries in most developed economies (Salamon et al., 2013).

A high voluntary workforce participation rate is a salient feature in the nonprofit sector. In New Zealand, of the 10.6% workforce representing the nonprofit sector, 6.2% represent the volunteer workforce while 4.4% represent the paid workforce; in Australia, of the 11.5% workforce representing the nonprofit sector, 3.0% represent the volunteer workforce while 8.5% represent the paid workforce (Salamon et al., 2013).⁵

In addition to its size and the sizable volunteer workforce participation rate, the Australasian nonprofit sector also stands out on two other characteristics: the extensive service focus and the revenue structure. In Australasia, the key service functions healthcare, education, social services and housing absorb more than half of the nonprofit sector workforce, paid staff and volunteers combined (Sanders et al., 2008). These service activities account for 62% of the gross value added (GVA), relative to the so-called “expressive activities” such as culture, sports and recreation, environmental protection, civic activism, labour unions and professional associations, and religious worship (Sanders et al., 2008). As regards the revenue, Australasian NPOs receive a

⁵ According to Salamon et al. (2013), who examined the characteristics of the nonprofit sector in 13 countries, New Zealand and Norway happened to have the highest proportion of volunteer workforce.

substantial portion of its income from fees (NZ – 67% and Australia – 51%) charged for the provision of goods and services rather than funding from the government or through philanthropy (Salamon et al., 2013). However, the Australian and New Zealand governments provide substantial contribution for the healthcare (nonprofit) sector in their respective countries (Crampton, Lay-Yee, & Davis, 2004; Scaife et al., 2015). Regarding philanthropy in Australia (better known as ‘giving’ in Australasia), Scaife et al. (2015) observe that although Australians are extremely generous when private donations are solicited for disaster-related incidents (they cite the 2004 Asian Tsunami as an example), philanthropy/giving is something that has not been widely embraced by the Australian society still, although there are signs of more philanthropic uptake.

As regards the cost structure, Australian NPOs devote majority (71%) of their expenditure on labour costs rather than on intermediate consumption (i.e. purchase of goods and services from other industries for use) in producing output, while New Zealand NPOs devote more (54%) on intermediate consumption rather than on labour costs (Salamon et al., 2013). New Zealand’s high intermediate consumption has been attributed to several factors, such as higher costs associated with the goods and services purchased by NPOs to deliver their services, a greater support from independent contractors or volunteers (this lowers the labour cost outlays), and low wage rates paid to nonprofit workers (Sanders et al., 2008).

2.5. CLASSIFICATION OF HEALTHCARE ORGANISATIONS

The organisations that provide healthcare services (general and special), advice on health related issues on nonprofit basis that are neither private nor public are categorised as nonprofit healthcare organisations. In the ICNPO, these organisations are further divided into 4 main subgroups according to the services they provide, as shown in Table 2.2.

Table 2.2: The ICPNO Sub-classification of the Health Sector (Salamon & Anheier, 1997)

Subgroup	Units	Description
Hospitals and Rehabilitation	Hospitals	Preliminary places that provide medical attention or treatments to illnesses or injuries
	Rehabilitation	Inpatient, outpatient and community medical services to people with physical damages, familial imperfections or disease and necessitating physio therapy or similar kind of attention
Nursing Homes	Nursing Homes	Inpatient recovering care, preliminary medical care services and homes for elderly people also included
Mental Health and Crisis Prevention	Psychiatric hospitals	Inpatient medical care and medical procedures or applications that are intended to relieve mental disabilities
	Mental Health Treatment	Outpatient medical procedures or applications that are intended to relieve mental disabilities, community mental health centres also included
	Crisis Intervention	Outpatient and counselling services for crisis situations, suicide prevention and helping abused people also included
Other Health Services	Public health and wellness education	Health educational and promoting programmes, first aid training and family planning programmes also included
	Health treatment, primary outpatient	Health clinics and vaccination centres are included
	Rehabilitative medical services	Includes nature cure centres, yoga centres and physical therapy centres
	Emergency medical services	Immediate care services such as ambulance services

2.5.1. Classification of Australasian Nonprofit Healthcare Organisations

Salamon and Anheier (1997) assert that because the function of the nonprofit sector in any country is ultimately formed by the entire pattern of social, economic and political development of that country, a country can adapt the ICNPO sub-classification to suit its context. Table 2.3 and Table 2.4 depict the classification of Australian and New Zealand nonprofit healthcare sectors respectively. It is clear that the degree of similarity between the Australian and the New Zealand classification systems is very high, which is a very important factor in this study, because the operational definitions of the PM dimensions are expected to hold true for both Australian and New Zealand healthcare NPOs.

Table 2.3: Classification of the Australian Nonprofit Healthcare Sector (Australian Bureau of Statistics, 2015)

Subsector	Description
Hospitals and rehabilitation	Primarily inpatient medical care and treatment and inpatient health care and rehabilitative therapy to individuals suffering from physical impairments due to injury, genetic defect or disease and requiring extensive physiotherapy or similar forms of care.
Nursing homes	Inpatient convalescent care and residential care, as well as primary health care services; includes homes for the frail elderly and nursing homes for the severely handicapped.
Mental health and crisis intervention	Inpatient care and treatment for the mentally ill and outpatient treatment for mentally ill patients; includes community mental health centres and halfway homes and outpatient services and counsel in acute mental health situations; includes suicide prevention and support to victims of assault and abuse.
Other health services	Public health and wellness education, rehabilitative medical services and emergency medical services includes ambulatory services and paramedical emergency care, shock/trauma programs, lifeline programs and ambulance services.

Table 2.4: Classification of the New Zealand Nonprofit Healthcare Sector (Sanders et al., 2008)

Subsector	Description
Hospitals and rehabilitation	Hospitals providing inpatient healthcare including physiotherapy and other rehabilitative therapy for those suffering from injury, genetic defect or disease.
Nursing homes	Inpatient convalescent care and residential care, nursing homes for the severely handicapped, hospice services.
Mental health and crisis intervention	Psychiatric hospitals, outpatient treatment for mentally ill and outpatient services and counselling in acute mental health situations.

Subsector	Description
Other health services	Public health promotion and health education, outpatient health treatment, outpatient rehabilitative medical services and emergency medical services.
Support and ancillary services	Nonprofit institutions not adequately covered in the activity descriptors above but related in purpose.

2.6. PERFORMANCE MEASUREMENT CHALLENGES IN NPOS

The major difference between NPOs and for-profit organisations is that the former have goals and services that are more intangible and harder to define (Kanter & Summers, 1994), which means that PM in NPOs is inherently challenging (Forbes, 1998; O'Neill & Young, 1988). Donors, people receiving the services (beneficiaries or clients) and the workers have different views on service outcomes (Bryson, 2011). In PM, this translates to measuring social value creation, rather than financial value creation (the present value of future returns). Stated alternatively, NPOs are mission-driven rather than being profit driven (Kaplan, 2002). Allocation of resources within for-profit organisations becomes more flexible in a number of ways, if they meet financial performance goals. NPOs on the other hand face limitations on their flexibility to follow suit, due to difficulty of achieving the desired social values (Kanter & Summers, 1994). For a mission-driven organisation, measuring “success” is very difficult (Bryson, 2011; Forbes, 1998; Kanter & Summers, 1994). Also, it appears that the technical difficulties of measuring complex missions or the political difficulties of designing measurement systems that accommodate the various stakeholders in NPOs has dampened the level of ‘mission measurement’ within the sector (Ritchie & Kolodinsky, 2003). This is not the case for the for-profit sector because sufficient freedom is available for these organisations to redefine their missions so as to make profit. Also, in a free enterprise system devoid of entry and exit barriers, for-profit organisations can move in and out of the business at any time, depending on the conditions of the external environment (Porter, 2008). However, NPOs have to stick to their missions to sustain their existence (Kanter & Summers, 1994).

Due to these well-document issues, the complexity of measuring performance in the nonprofit sector has long been recognised (Moxham, 2009; Sawhill & Williamson,

2001). While for-profit sector organisations successfully use proven conventional business models to measure performance in the areas of strategic planning, marketing, finance, information systems and organisational development, the NPOs use the same models in the hope that the use of these models will benefit them the same way as they benefit for-profit organisations (Moore, 2003). Since the for-profit sector's success is mainly associated with profit and loss accounts figures and availability of many quantitative data to measure profits, the same performance approaches do not create the same results for NPOs (Kanter & Summers, 1994; Moore, 2003).

Studying organisational effectiveness in NPOs through standard PM tools also becomes complex due to their legal and financial status. NPOs also often pursue unstructured goals pertaining to delivery of intangible services, which limits the use of quantitative measures used in the for-profit sector. Finally, as Kanter and Summers (1994) note, the strategic objectives of NPOs are often based on “societal values about which there may be little or no consensus”, which makes PM even more challenging. For these reasons, the researcher argues that NPOs need to develop measurement systems that serve their own strategic needs.

2.7. CHAPTER CONCLUSION

This chapter began with an overview of NPOs. The chapter also provided an operational definition of NPOs. This definition reliably distinguishes NPOs from for-profit and government organisations, which in effect, justified why NPOs need separate PM systems. The complementary concepts ‘governance’ and ‘accountability’ were also reviewed, particularly in relation to NPOs. Knowledge of nonprofit governance and accountability is important to the researcher in conducting her fieldwork (the case study component of the study). In addition, the theoretical perspectives that the researcher reviewed on governance (particularly agency theory and stakeholder theory) are relevant to the researcher's methodology and discussion. Chapter two also reviewed the nonprofit sector in Australasia and how the nonprofit healthcare sector is structured in the region. This provided important contextual information for the researcher's research design. The next chapter covers the core literature on PM systems, which enabled the researcher to identify the knowledge gaps to design research that makes a valuable contribution to the existing body of knowledge on PM systems.

CHAPTER THREE

CORE LITERATURE ON PERFORMANCE MEASUREMENT SYSTEMS

Success depends upon previous preparation, and without such preparation there is sure to be failure

Confucius

3.1. INTRODUCTION

This chapter is the first of the two chapters that review the literature important to the research. The contextual literature relevant to the study and the knowledge gaps identified in the literature as well as the research questions are presented in this chapter.

Performance measurement (PM) refers to the process of collecting, analysing and reporting information pertaining to the performance of any entity: an individual, a social entity such as an organisation, a physical object, or a physical system (LeRoux & Wright, 2010; Lynch-Cerullo & Cooney, 2011). However, in this study, the focus is solely on “Organisational Performance Measurement”—that is PM at business/organisational level (Neely, 2007). Therefore the reader should note that from this point onwards, the researcher uses the term PM to mean Organisational Performance Measurement⁶ only.

This chapter is organised as follows. Section 3.2 covers the literature that justifies PM systems in organisations and how PM systems have evolved since their inception at the turn of the twentieth century. Section 3.3, one of the main sections of the chapter reviews the literature on key PM frameworks: the Balanced Scorecard (BSC), Performance Prism (PP), and two Performance Excellence Frameworks; in this section a particular emphasis has been given to examining the conceptual foundations of the aforesaid PM frameworks through critical synthesis of the literature. The BSC, by far

⁶ Organisational level PM is sometimes called “Business Performance Measurement” in the case of for-profit organisations (Neely, 2007).

the most scrutinised PM framework in the literature gets more coverage in this section. The next section (section 3.4) reviews PM frameworks in the context of nonprofit organisations (NPOs), based on the available literature. Therefore literature on two additional PM frameworks is also reviewed in section 3.4, although again, the BSC occupies more coverage, due to its relevance to the researcher's study. Section 3.5 reviews prior studies (implementation studies, conceptual papers, and empirical studies) that examined different facts of the validity of the BSC in a nonprofit context. Section 3.6 identifies the knowledge gaps based on the researcher's literature synthesis, along with the corresponding research questions. Finally, section 3.7 concludes this chapter through a concise summary of key points pertaining to the major sections of this chapter.

3.2. THE RATIONALE FOR PERFORMANCE MEASUREMENT AND ITS EVOLUTION

Oakland (2001) showed that PM as applied to processes and organisations is best understood by analysing the meanings of the words “*performance*” and “*measurement*”. He defined performance as the results on organisational outputs⁷, their outcomes⁸, as well as the processes (a series of interconnected processes produce outputs through tangible and intangible inputs) that create the outputs and the inputs (resources) to go into the processes. Therefore Oakland takes a “systems perspective” (Daft & Weick, 1984; Kast & Rosenzweig, 1972) on organisations in assessing their performance, in that performance is not solely about outputs and outcomes but also on what happens within the system (organisation) and what goes into it. Oakland (2001) and others (e.g. Adebajo, Abbas, & Mann, 2010; Watson, 1993) show that the products and services that an organisation produces (outputs) as well as its processes and inputs permit evaluation and comparison relative to the goals, standards, past results, and what happens in other organisations. Performance can be expressed in both financial and nonfinancial terms (Kaplan & Norton, 1992; Neely, Marr, Roos, Pike, & Gupta, 2003). Measurement refers to gathering and analysing numerical information that quantifies input, output, and performance dimensions of inputs, processes, products, services, and the overall organisation (Neely et al., 2005; Oakland, 2001).

⁷ An example might be the number of customers served within a calendar month.

⁸ An example might be revenue (in a case of a for-profit organisation) and/or social impact. A social impact can be as a result of a commercial activity or a nonprofit/charitable activity.

In keeping with the definition of Neely et al. (2005) on a PM framework, PM can be defined as a process of defining, monitoring, and using objective indicators of performance of the organisation and its programmes on a regular basis. It involves creating a simple, but an effective and efficient system for determining whether or not an organisation meets its objectives (LeRoux & Wright, 2010; Lynch-Cerullo & Cooney, 2011; MacIndoe & Barman, 2013). All processes warranting measurement (for performance evaluation) require the use of statistical information to determine the effectiveness of the process. A full scope of the performance of an organisation can never be captured, as generally, some performance parameters cannot be measured directly. Instead, performance must be estimated via indirect observations that are compressed into a set of numerical measures known as performance measures. According to the literature, the management information fed through suitable performance measures inform whether or not the senior management of the organisation has implemented the long range (strategic) plans effectively by allocating the organisation's scarce resources effectively (LeRoux & Wright, 2010; Lynch-Cerullo & Cooney, 2011; MacIndoe & Barman, 2013). Thus PM is an important strategic activity undertaken by an organisation towards improvement—that is to guide progressing towards achieving the targets (strategic goals).

Literature reveals that PM is vital to an organisation (nonprofit or otherwise), due to the following reasons⁹:

- PM is intended to produce objective and relevant information on a specific project or programme (e.g. health related advice for underprivileged school children) or the overall organisational performance. Measuring the overall organisational performance helps organisations in: (a) strengthening the existing management systems, (b) high impact decision making, (c) achieving results and (d) improving the overall performance (LeRoux & Wright, 2010). In a nutshell these activities are referred to as (strategic) *performance management* (Otley, 1999; Schläfke, Silvi, & Möller, 2013).

⁹ The reasons are not mutually exclusive.

- PM tends to focus on what is being measured and on the performance itself; PM motivates people and organisations to work on improving the performance (LeRoux & Wright, 2010).
- PM helps managers and other employees to assess the state of the performance of their organisations and determine the capacity requirements to plan and implement effective improvement programmes (LeRoux & Wright, 2010; MacIndoe & Barman, 2013).
- PM results in identifying opportunities for process improvement (Evans, 2004; Moore, 2003)
- PM provides a clear focus on mission and strategy; PM improves management decision making, improves the organisational performance itself and increases the management accountability to governing bodies and stakeholders (Lynch-Cerullo & Cooney, 2011; Poister, 2003).
- Given that a proper management system is required to achieve targets, PM supports management functions such as strategic planning, budgeting and financing, programme management and evaluation, quality and process improvement, contract management, external benchmarking, and communication with other parties (Kanter & Summers, 1994).
- PM helps to use resources more effectively by being able to exercise more control over operations by allowing more flexibility at operating levels. Moreover, it helps to evaluate the efficiency of alternative approaches (Kanter & Summers, 1994).
- Throughout the last 20 years, both funding bodies and the public have increased pressure for accountability and performance of the corporations (NPOs included), necessitating them to use PM systems (LeRoux & Wright, 2010; Poole, Nelson, Carnahan, Chepenik, & Tubiak, 2000). Measuring performance is the one way of re-establishing the *public confidence* in NPOs (Carman, 2009).

- The phenomenal growth of the sector, more professionalism, and the expansion of the stakeholder base in the nonprofit sector have highlighted the importance of accountability (Benjamin, 2013; LeRoux & Wright, 2010).

Although there were various PM practices in use even in the industrial era, the first comprehensive implementation of a PM system originated around 1906 at the New York Bureau of Municipal Research (Williams, 2004). Over the next thirty years, these practices became more advanced through increased quantification and based on past experience of the experts. Early PM practices fulfilled two objectives. First, they communicated information to managers to show the nature and the state of work completed, which contributed towards improved work productivity through management intervention. Second, they supported budgetary decision making by divulging information about the budgetary requirements and the effectiveness of budgetary allocations (Williams, 2003). In essence, older PM systems focused solely on the financial dimension of the business enterprise; since the 1980s this met with criticism as being a myopic view of the performance of a firm (Neely et al., 2003). As a consequence, systems for assessing both financial performance and non-financial performance were developed in the late 1980s and early 1990s (Neely et al., 2003). One of the most well-known systems is the BSC (Kaplan & Norton, 1992), which is now widespread in organisations belonging to all three sectors: private, public, and nonprofit. Other systems, such as the SMART pyramid (Cross & Lynch, 1988), and the PP (Neely et al., 2001) have also been developed and promoted.

The PM systems generated in the above era (late 1980s and early 1990s), also known as *first generation PM frameworks*, provide guidance to the managers which non-financial measures they should adopt in their organisation's PM system. Neely et al. (2003) observe that the first generation systems focused on generating performance metrics and not so much on the relationships between individual performance measures. They also show that *second generation PM frameworks* were developed as a solution for the problem that senior managers face as to how individual measures should be integrated as a coherent strategic performance monitoring system. According to them, second generation PM frameworks made a significant step forward as they started to address the problem by mapping the causal flows and transformations. The current systems, *third generation PM systems*, move even further forward by attempting to link the non-

financial and intangible dimensions of organisational performance and the cash flow effects in a more coherent manner (Neely et al., 2003). In spite of all these improvements, the researcher identifies some knowledge gaps, which are reviewed in this chapter.

3.3. PERFORMANCE MEASUREMENT FRAMEWORKS

A PM framework/system can be defined as a “set of metrics used to quantify both the efficiency and effectiveness of actions” implemented to achieve the strategic goals of an organisation (Neely et al., 2005, p.1229). The theoretical framework of a PM system facilitates the generation of performance metrics along predefined measurement dimensions. For each PM dimension, managers need to determine the appropriate set of measures to gauge the organisational performance (Kennerley & Neely, 2002) and the measures themselves, should possess the following features: (a) they should be derived from the strategy (Anthony & Govindarajan, 2007; Neely et al.,2005); (b) they, as a whole should be capable of providing a balanced view of performance (Kaplan & Norton, 2001; Neely, 2007); and (c) encourage appropriate actions for goal achievement (Kaplan & Norton, 2001; Neely, 2007). The Balanced Scorecard, the Performance Prism, and national quality award frameworks (especially the Baldrige Framework originated in the USA and the EFQM Excellence Model originated in Europe) are some of the well-known PM frameworks that are being used by managers to develop appropriate performance measures to suit the strategic objectives of their organisations to monitor and improve strategic performance (Hoque, 2014; Perkins, Grey, & Remmers, 2014). In terms of scholarly attention, the researcher observes that the BSC has received considerably more attention than the other PM frameworks, over the past 15 years.

Figure 3.1 depicts the number of publications on four selected PM frameworks (models).¹⁰

¹⁰

Figure 3.1 was created by collecting data via the search engine “Scopus”, using appropriate search words in the fields “Article Title”, “Abstract”, and “Key Words”.

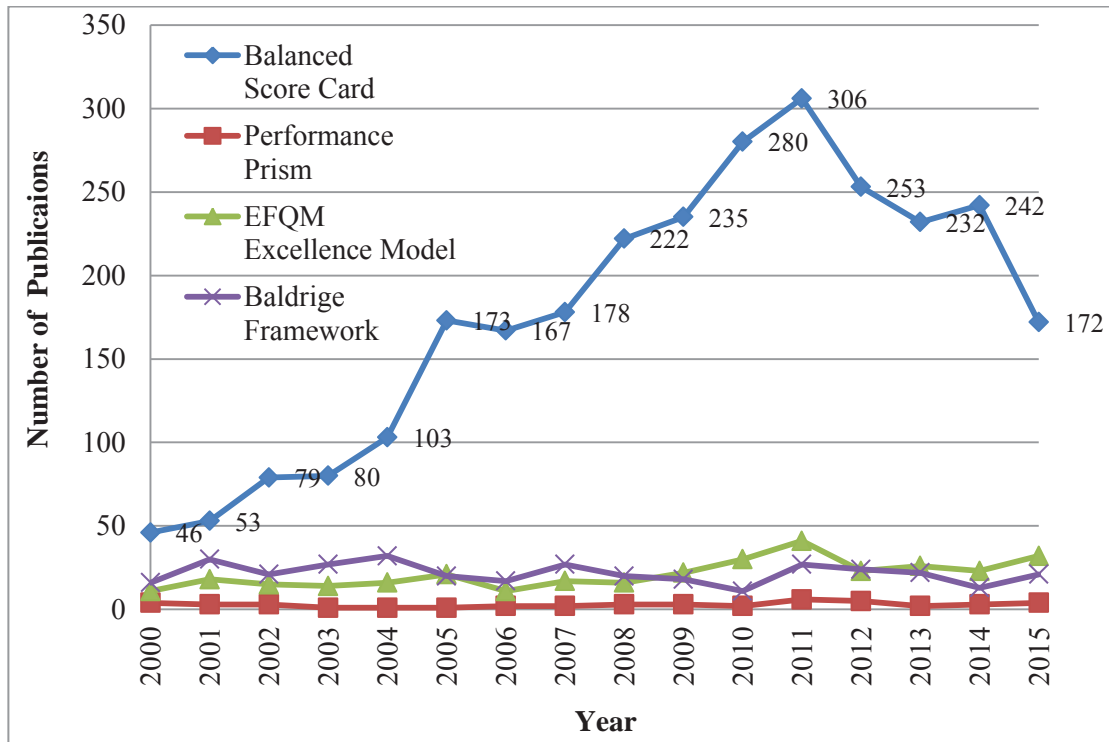


Figure 3.1: Number of publications on four selected PM frameworks

3.3.1. Role of PM Frameworks and Empirical Support for Their Effectiveness

Otley (2001) argues that the accounting community implicitly recognise that PM systems have three fundamental roles to play in organisations. Firstly, a PM system provides a tool for financial management (i.e. a tool that facilitates efficient and effective management of funds to achieve the strategic objectives of the organisation); secondly, it provides information on the overall business performance; and thirdly, it provides a means of motivation to perform, exercising control at the same time. Wholey (1999) emphasises that *quality of information*, *utility for decision making*, and *cost effectiveness* are the most vital features of PM systems. For these reasons, it is very important for an organisation to adopt a suitable PM system.

Any research that develops new PM models or works on improving the existing PM models (such as this research) would be of little value if the literature does not support the proposition that PM leads to improved overall organisational performance. The literature supports the proposition that PM is positively related to organisational performance both in the for-profit sector (e.g. Tung, Baird, & Schoch, 2011) and more

importantly, in the nonprofit sector (e.g. Francisco & Alves, 2012; LeRoux & Wright, 2010; Limburg, Knowles, & McCulloch, 2012).

Based on large sample survey data from the US NPOs, LeRoux and Wright (2010) empirically studied the strength of the relationship between the dependence of performance measures and actual strategic performance. They found that there is a significant positive relationship between the breadth of performance measures used (i.e. the number of performance measures and the extent to which they span across multiple dimensions) and strategic performance. They also found that there were also other factors that contribute towards effective strategic performance: “governance”, diversity of funding and the “education level of the executive directors”. Limburg et al. (2012) examined the literature to show that whilst PM is positively related to strategic performance in a NPO, the relationship is not as strong as in a for-profit organisation. They showed that the main reasons for this are: having to have multiple purposes (e.g. NPO has to achieve its mission but it must also ensure financial stability), complex stakeholder environment, and having to report to multiple stakeholders who need different information from the NPO. They argued that a PM system of a NPO can improve its effectiveness through an “enterprise performance management system”—that is to manage a NPO much like a business enterprise that creates value for its stakeholders. Unfortunately, the authors do not explain how this could be accomplished, given the complex issues raised by themselves, such as the complex stakeholder environment and divergent strategic objectives.

Using a single Portuguese case study (in-depth interview of managers and scrutinising of administrative documents), Francisco and Alves (2012) investigated how multiple stakeholder interests can be incorporated into a PM system to make easier assessments of the organisational performance. They found that whilst the NPO that they studied used several measures, the managers were concerned about not having a simple and fundamental performance measure such as the “profitability” as used in a for-profit organisation. Francisco and Alves proposed a multidimensional framework based on the “stakeholder theory” (Donaldson & Preston, 1995; Freeman, 1984) to improve the effectiveness of the PM system of a NPO.

MacIndoe and Barman (2013) assert that NPOs do not operate in a vacuum in that they are influenced by many different parties (stakeholders). They show that the managers of NPOs have to adopt suitable PM systems to operate their organisations, taking into account the complex environment in which they operate; as examples of environmental complexity, MacIndoe and Barman cite different funding sources, increased demand for accountability from the public and the donors as well as competition for donor funding from other NPOs.

While measuring the performance of NPOs is vital (Alexander, Brudney, & Yang, 2010), literature shows that it is more difficult to measure performance of NPOs than for-profit organisations (Bryson, 2011; Duque-Zuluaga & Schneider, 2008; Forbes, 1998; Kanter & Summers, 1994; Soysa et al., 2013a). Several reasons for this has been cited: different stakeholder expectations, difficulties in aligning mission and programmes, difficulties in defining organisational *outputs* and *outcomes*, difficulties in using financial terms to measure organisational success (or failure), limitation on the use of quantitative measures due to pursuing unstructured goals for delivery of intangible services (Benjamin, 2013; Duque-Zuluaga & Schneider, 2008; Kanter & Summers, 1994; MacIndoe & Barman, 2013; Soysa et al., 2013a). It is often argued that the bottom line of a NPO is “social profit” (Bouckaert & Vandenhove, 1998; Candler & Dumont, 2010; Cutt, 1982; Schneider, 2009). Bouckaert and Vandenhove (1998, p. 1075) operationalise social profit as “cost-efficient use of the available capital such that the *social objective, as defined by the organisational authority*, is achieved to its best”.

The researcher notes that a number of PM frameworks (models) have been prescribed in the literature for measurement of performance of projects and organisations concerning social impact. Maas and Liket (2011) reviewed 30 such frameworks to identify their similarities and differences in terms of the approach and other criteria. They observe that PM frameworks can fall into three types, in terms of their approach: process focused models, outcome focused models, and monetarisation models. Maas and Liket contend that because process models primarily focus on the metrics of efficiency and effectiveness of the organisation’s operational processes, inputs, and outputs, they do not provide specific metrics on social returns at the absolute level. They nonetheless observe that process models are useful because managers can select process output

metrics that correlate with the financial, social, and environmental outcomes. The researcher observes that PM models found in the quality and operations management literature (for both for-profit and NPOs) are invariably process-based models. The conceptual foundations of 4 PM frameworks relevant to the researcher's study are reviewed next. These are process-based PM frameworks that can be used to monitor the strategic performance of the whole NPO.

3.3.2. The Conceptual Foundation of the BSC

The BSC was developed by Kaplan and Norton (1992) as a management accounting tool (Kaplan was a professor in accounting at the Harvard Business School at the time of publishing the BSC in 1992) for private corporations (Kaplan, 2008; Kaplan & Norton, 1992). Kaplan and Norton (1992) premised that private corporations are “profit driven” in that they exist to maximise value to the shareholders. However, they showed that traditional financial measures alone do not provide sufficient management information for the managers who manage the organisation on behalf of the shareholders (Kaplan, 2008). Kaplan and Norton showed that value to shareholders (profitability) result in satisfying the external customers (who buy the goods or services produced by the corporation) and that customer satisfaction results from managing the processes efficiently and effectively by using the organisational resources—in particular, people and knowledge. Therefore Kaplan and Norton asserted that management reporting should strike the right balance between information on financial performance and information on the antecedents of financial performance (Kaplan, 2008). Consequently, the BSC entails four measurement perspectives (dimensions) aligned in a sequential order: Learning and Growth, Operational (processes), Customer, and Financial (Kaplan, 2008; Neely, 2007).

Figure 3.2 depicts the original BSC framework published by Kaplan and Norton (1992), based on one year-long research that involved studying the PM systems of 12 companies who excelled in performance monitoring. The purpose of the original BSC framework was to help companies in generating financial and non-financial performance measures for the key elements of their organisational strategy (the strategy being expressed via strategic goals). According to Kaplan and Norton, the original BSC framework is based on the following propositions: (a) a company will look good to its

shareholders in the future only if it looks good to the customers at present, in terms of strategic goals on time (e.g. on time delivery in the case of existing products), quality, performance, and service; (b) a company will look good to its customers only if it excels at its internal business processes in meeting customer expectations; and (c) a company can excel at (and sustain) its internal business processes only if it engages in innovation and learning.

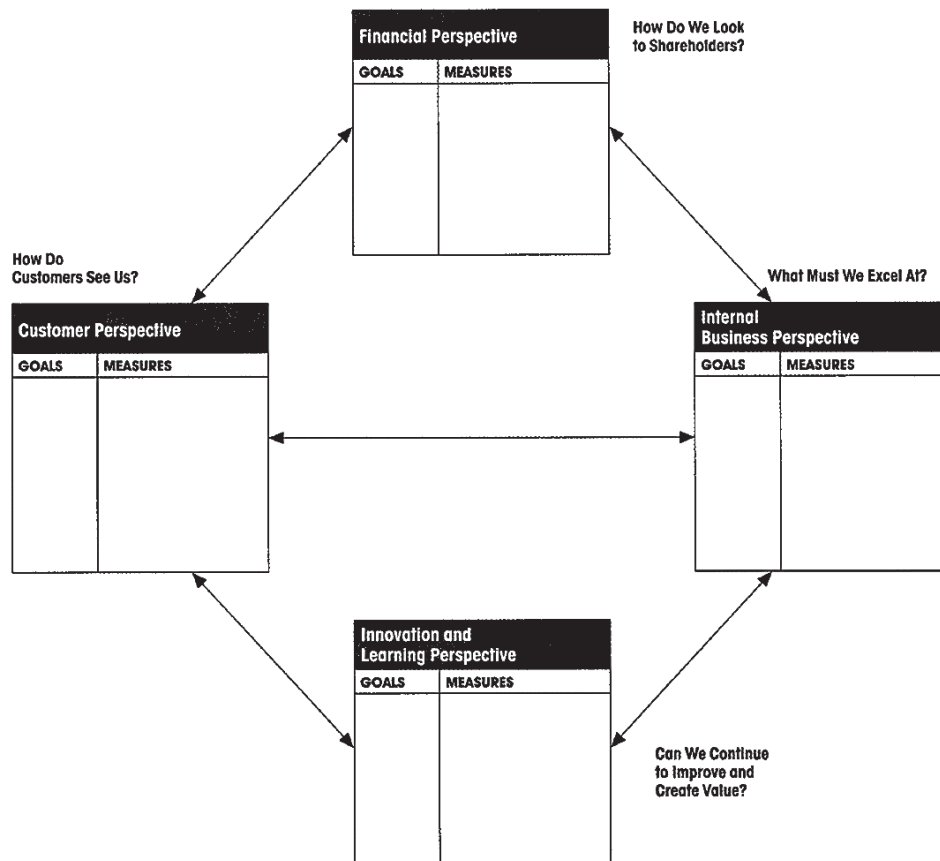


Figure 3.2: The original BSC framework published by Kaplan and Norton (1992)

The BSC has come a long way from being a mere management accounting tool to a strategic decision making mechanism through a number of revisions and refinements (Anon, 2008; Kaplan, 2008; Kaplan & Norton, 2001; Lawrie & Cobbold, 2004). It is generally agreed that the BSC has seen two major revisions (the four measurement perspectives have remained intact although the labels of some have changed) (Anon, 2008; Kaplan, 2008; Lawrie & Cobbold, 2004).¹¹ The first major revision, generally

¹¹ In addition, the BSC has been perfunctorily revised for public and NPOs, which remains the central focus in this study (the nonprofit BSC is reviewed later).

referred to as the second generation BSC (Lawrie & Cobbold, 2004), was published in 1996 (Kaplan & Norton, 1996a). The primary aim of this revision according to Kaplan and Norton (1996a) was to provide more guidance to managers as to how performance measures should be chosen (Kaplan, 2008; Lawrie & Cobbold, 2004). Figure 3.3 depicts the second generation BSC framework.

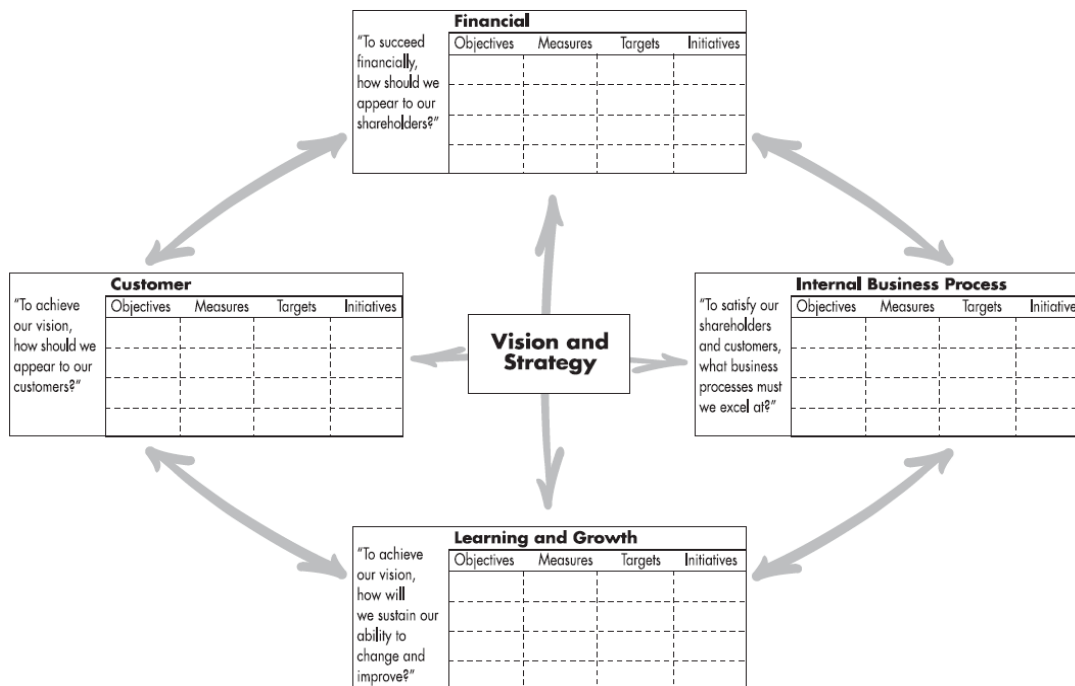


Figure 3.3: The second generation BSC framework (Source: Kaplan & Norton, 1996b)

The second generation BSC is based on the following premises: the vision and strategy of a business should be translated to the four PM perspectives (dimensions). Each PM dimension should contain multiple performance measures, each measure corresponding to a specific strategic objective. There should be a causal connection between performance measures, and the performance dimensions themselves are causally related in the following causal order: learning and growth → process and customer → financial (the subsequent versions of the BSC provided more clarity on the causal order).

The second revision (this took place in late 1990s, according to Lawrie & Cobbold, 2004), which sees the BSC in its current form, typically known as the *third generation BSC*, retains the second generation BSC framework (Figure 3.3) but addresses causality between performance measures (also somewhat loosely, between the four PM

perspectives/dimensions) through an accessory known as the “strategy map”, which maps the linkages between strategic objectives (see Figure 3.4 as an example) to explain how the shareholder value is maximised. Therefore, it can be argued that the founder/s of the BSC made incremental improvements to the BSC to align performance measurement to strategy planning and monitoring.

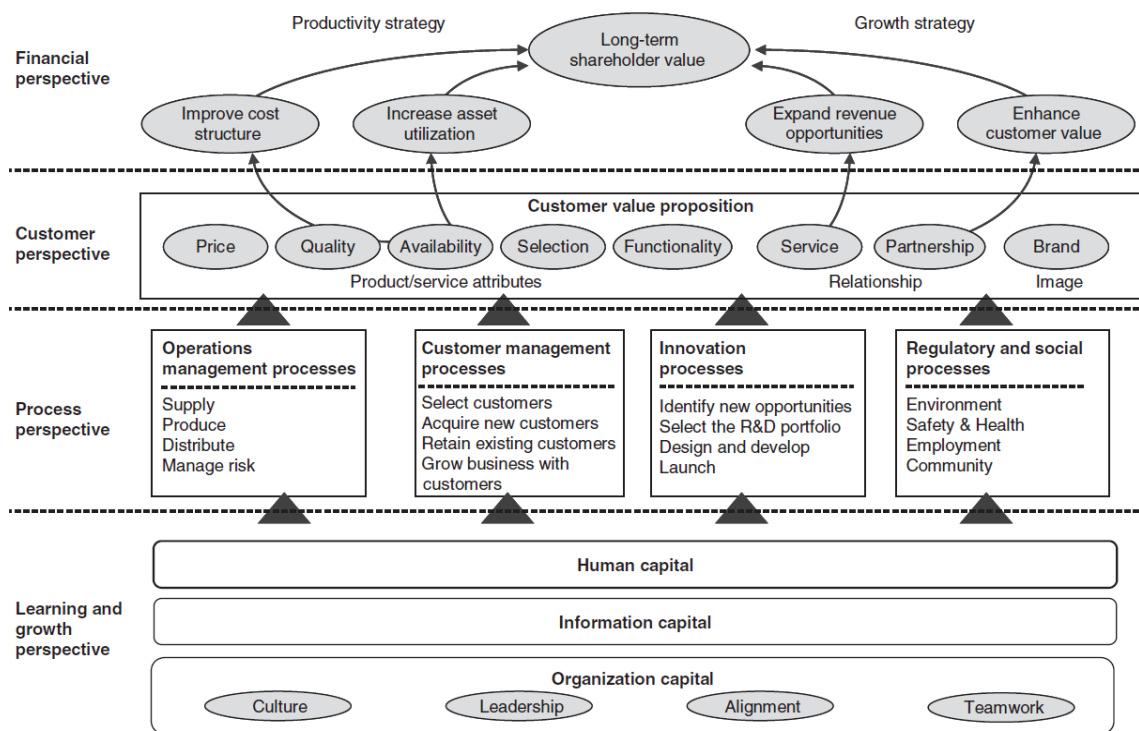


Figure 3.4: A typical strategy map of a for-profit organisation embodied in a third generation BSC (Source: Kaplan, 2008)

Another very important development with regard to the BSC is the publication of a separate paper by its co-founder Kaplan (2008) to articulate the conceptual foundations of the BSC (the third generation BSC to be more specific). This article provides answers to some of the misconceptions about the BSC. Kaplan (2008) argues that although the BSC appreciates the fact that the stakeholder needs and their interplay (the needs of the shareholders, customers, communities, employees and suppliers) is central to strategy planning, the BSC does not assume that strategy formulation begins by identifying the needs of the stakeholders, as advocated by *stakeholder theorists* (e.g. Freeman, 1984). This is because—according to Kaplan—the stakeholder view provides a less powerful view of value creation of a firm, as joint maximisation of stakeholder

needs resulting in suboptimal long-term value of the firm (more precisely value to the owners). He argues that BSC is based on the assumption that formulation of the organisational strategy should precede stakeholder satisfaction and that the objective of a business is to maximise its value (strategies should be formulated to maximise value of a firm), given the complex interplay between the stakeholders. Similarly, Kaplan argues that the BSC inherently appreciates the importance of fulfilling the needs of the employees (the learning and growth perspective, among other things, takes into account intrinsic and extrinsic needs of the employees) without making employee satisfaction central to the organisation. The BSC assumes that employee satisfaction is only a means to an end and not the end (company value maximisation) in itself.

In the aforesaid paper, using the “agency theory” (section 2.3.1.1) as a frame of reference, Kaplan (2008) brings the self-interested behaviour of top managers into the limelight. Agency theory holds that there is often a conflict of interest between the interests of a firm’s shareholders (maximising the net benefits of future returns) and the interests of senior managers (maximising their own self-interests), who are the agents of the shareholders. Kaplan concedes that the BSC fails to provide a single metric that captures the overall performance of an organisation (performance measures in the BSC are unweighted) to enable the owners to devise an equitable reward system for the executives. However, according to Kaplan, linking firm performance to executive pay is only one of the many aspects of a good PM system. As a corollary to the self-interested agent proposition (agency theory), the researcher observes the BSC does not directly indicate how well organisational leadership is driving the organisation to achieve shareholder objectives (long-term financial performance). Four different PM perspectives and multiple measures (in each perspective) mean that in the absence of guidance as to how each PM perspective and its elements should be weighed, it becomes difficult to quantitatively represent the overall organisational performance.

Based on the above arguments of Kaplan (2008), who is a management accounting scholar as mentioned earlier, the researcher is of the view that the conceptual underpinnings of the BSC is heavily shaped by principles of managerial economics such as agency theory and the theory of the firm. In simple terms, the theory of the firm views the objective of a firm as maximising the net benefits of future returns (i.e. maximising long term shareholder value) (Brigham & Pappas, 1976; Jensen, 2002). In

spite of over 20 years of academic scrutiny, three conceptual revisions and evidence of widespread practical usage, the BSC (even for a profit making enterprise) is still not without its critics (Awadallah & Allam, 2015). Nørreklit (2000) and Nørreklit, Nørreklit, Mitchell, and Bjørnenak (2012) critiqued the BSC on two grounds. First, Nørreklit examined the implied causal relationships between the measures, which form the basis of the BSC (e.g. strategy map): “measures of organisational learning and growth → measures of internal business processes → measures of customer perspective → financial measures” (Nørreklit, 2000, p.68). Nørreklit et al. (2012) argued that the implied causal relationships between the measures in a BSC is incomplete, because financial performance cannot be reliably predicted this way (they argued that more variables are needed to predict financial performance reliably). Second, Nørreklit et al. examined the claim made by the founders of the BSC that BSC is not only a strategy monitoring system but also a strategic control system (i.e. the BSC can also be used to effect corrective action). Nørreklit (2000) and Nørreklit et al. (2012) contend that the BSC cannot be an effective strategic control system because: (a) it does not account for all the key external stakeholder relationships (they show that the BSC omits suppliers and public authorities), and (b) the top-down approach advocated in formulating the BSC does not foster the requisite internal stakeholder relationships (manager-employee relationships) for the BSC to get rooted in the culture and the management system of an organisation.

Other scholars also acknowledge deficiencies of the underlying theory of the BSC, the process of creation of the scorecard, and its rationale. Rillo (2004) asserts that the BSC does not consider certain important interest groups such as the government and employees. Maltz, Shenhar, and Reilly (2003) observe that the biggest shortcoming of the BSC is non-inclusion of employee satisfaction as a primary PM domain. They argue that without satisfied employees an organisation cannot achieve high levels of customer satisfaction and therefore financial performance. Pessanha and Prochnik (2006) criticise the BSC for overlooking stakeholder satisfaction in favour of shareholder satisfaction when selecting the strategic objectives and performance measures. Moreover, they contend that although the Internal Processes perspective and the Learning and Growth perspective capture different facets of workforce needs, the BSC does not measure actual level of employee satisfaction. What they imply is that an

organisation may take several initiatives to motivate staff intrinsically and extrinsically but there is no real guarantee that these initiatives actually result in satisfied employees. Parmenter (2012) criticised the BSC for lack of clarity (e.g. insufficient detail) in outlining how performance measures should be developed for each perspective. Parmenter compared the characteristics of key performance indicators (KPIs) yielded by a BSC with the characteristics of KPIs used by iconic organisations such as Toyota, Nissan, GE, and Honeywell (he called these “winning KPIs”) and identified inconsistencies in the BSC approach. He argues that the primary role of “winning KPIs” is to guide managers in focusing on “critical success factors” (CSFs) and that BSC’s performance measures on the other hand are static measures meant to monitor and evaluate the performance of strategies (the researcher views this as a convoluted argument). In the same vein, Parmenter argues that the BSC does not provide any direction as to how CSFs should be determined. Parmenter argues that CSFs are the key to identifying effective KPIs. Antonsen (2014) also contends that the BSC does not provide any guidance as to how CSFs could be identified, which is central for generating KPIs. Anand, Sahay, and Saha (2005) have also criticised the BSC for lack of guidance given in the KPI selection process.

Meyer (2002, p.184), an influential author on management and sociology, does not see any issue in using the BSC framework in the strategy planning and implementation, provided the “hidden assumptions” in the BSC are made explicit. To him, there are two major hidden assumptions in the BSC framework: (a) strategy formulation is a top-down approach and therefore, essentially a senior management prerogative, and (b) operational level performance measures can be intuitively deduced from strategic level performance measures that are derived from the strategy maps. Meyer also shows that the BSC does not indicate how the performance, as indicated by each individual measure, collectively explain the overall performance.

Drawing inferences from the *theory of “bounded rationality”* (Simon, 1957)¹², based on his research experience on experimental behavioural accounting, Salterio (2012) argued

¹² Simon’s theory of bounded rationality (Simon, 1957) posits that the human brain is bounded by its ability to process information. Salterio (2012) also refers to the magic number 7 ± 2 pieces of information together as the limit and go on to show that Kaplan and Norton’s prescription (Kaplan & Norton, 1992) of 4 to 7 measures per measurement perspective is consistent with the theory of bounded rationality.

that the BSC attempts to limit the number of performance domains it covers and the measures being used, at the expense of articulating a comprehensive theory on business success. Salterio also opined that in the actual implementation of the BSC, users use different subjective weights for different measures (more commonly agreed measures being given more weight) based on their perception of each measure, which also results in some dimensions receiving more weight over the others, in the strategic monitoring process. While this idea is generally consistent with the principles of multi-criteria decision making (e.g. the analytical hierarchy process), the researcher sees the opportunity of using an empirical model fitting approach to address this issue more scientifically (see section 3.6).

Dechow (2012) questioned the validity of the BSC's rhetoric "BSC translates strategy into action" through the concepts and objects embedded in the BSC. Notwithstanding the fact that the two founders of the BSC (Kaplan and Norton) have assured that it is the case, Dechow questioned whether practitioners ever use strategy maps (a key object in the third generation BSC) in strategic planning process. He argued that future research should examine the link between the BSC concepts and its objects (e.g. PM dimensions, strategy maps, and the casual links) more closely and also whether the BSC actually translates corporate strategy into operational action.

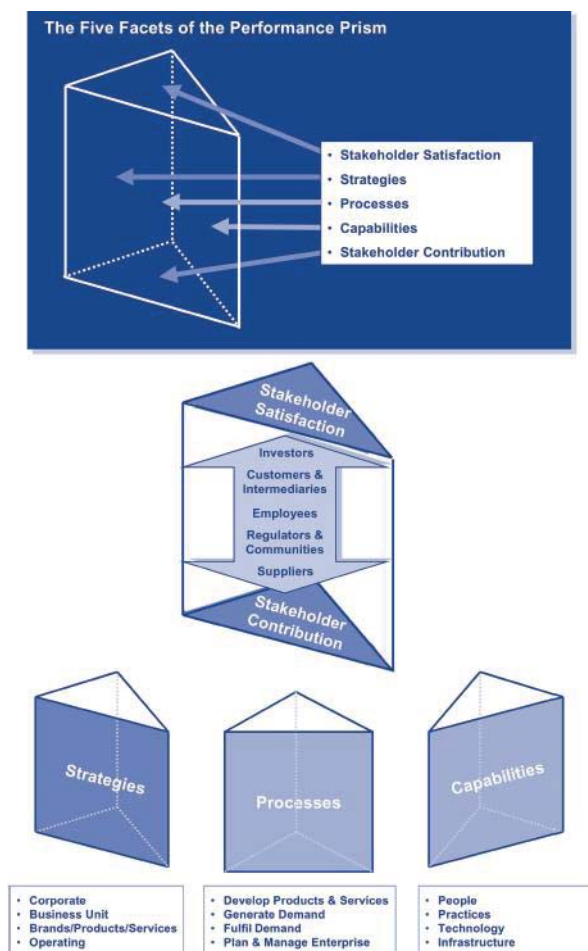
The researcher observes that some criticisms (e.g. lack of consideration of all the key stakeholders, non-inclusion of a dedicated measurement domain for employee satisfaction) levelled at the BSC are partially attributable to lack of full understanding of the conceptual foundations of the BSC, on the part of the critics. In fairness to some critics (those who criticised the BSC prior to 2008), Kaplan articulated the conceptual foundations of the BSC much later (in 2008, as discussed earlier), in part, to clear some misconceptions of the BSC. More recently, in a special issue on the BSC, Kaplan (2012) responded to commentaries made by critics (in the same issue) on the BSC. Some papers relate to complexities that arise in multi-stakeholder environments and BSC's apparent inability to account for these adequately, which is, in fact, not a new idea. Two other papers of relevance—Salterio (2012) and Dechow (2012)—have been covered in this section. However, the researcher sees merit in some criticisms of the BSC (more notably, over-emphasis of the shareholder paradigm and inadequate attention to employee satisfaction), in the case of application of the BSC in NPOs in its

current form. Being a focal area of this study, gaps in the BSC are further reviewed in sections 3.4 onwards (in relation to NPOs). In essence, the researcher supports the notion that the BSC requires more clarity, because in its current form, it requires significant alterations to fit to the specific requirements of each organisation, often involving the help of external consultants (Hoque, 2014; Qu & Cooper, 2011; Noell & Lund, 2002).

3.3.3. The Conceptual Foundation of the PP

In early 2000, Neely and his colleagues (Neely et al., 2001) designed a PM framework known as the *Performance Prism* (PP), using an accessory that was termed the “success map”; the success map, Figure 3.5 embraces the organisation strategy into the management system.

The Performance Prism Framework



An Example of a Success Map

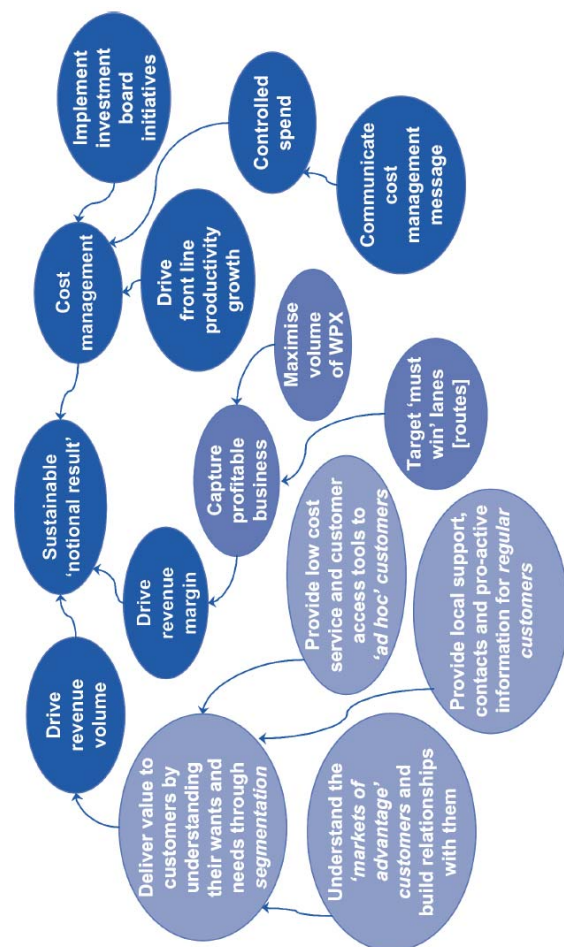


Figure 3.5: The PP framework and an example of a success map (Source: Neely, Adams, & Crowe, 2001)

The researcher observes that in the literature (e.g. Bourne, Kennerley, & Franco-Santos, 2005; Carlucci & Schiuma, 2007; Koufteros, Verghese, & Lucianetti, 2014) a success map is viewed as being equivalent to a strategy map. This is because success is the - outcome of the strategy and therefore causal dependencies among strategic objectives (BSC) should be consistent with causal dependencies among successes (PP). Five interrelated facets—Stakeholder Satisfaction, Strategies, Processes, Capabilities, and Stakeholder Contribution—are embodied in the PP (Figure 3.5). The first dimension Stakeholder Satisfaction provides a wider understanding on the stakeholders—employees, suppliers, partners, and the community—compared to the BSC and other PM frameworks (Neely, 2007). The second dimension Strategies, which follows from the first dimension, concentrates on organisational-level strategies. According to Neely (2007), the PP emphasises the importance of deriving strategies according to the stakeholders’ expectations, in line with the stakeholder theory.¹³ PP requires all key processes (the third dimension) that have been established to conduct organisational functions be considered in identifying performance measures under the Processes dimension. The researcher observes that conceptually, the Processes dimension in the PP is analogous to the Business Processes dimension in the BSC. The fourth dimension Capabilities captures the tangible and intangible assets of the organisation (employees, rules, skills, machinery and the organisational setup) required to drive the processes. The researcher observes that conceptually, the Capabilities dimension in the PP is analogous to the Learning and Growth dimension in the BSC. According to Neely (2007), the said four PM dimensions of the PP are the dimensions that strengthen an organisation’s competitive advantage (in a nonprofit context, the organisations still do compete with one another for resources such as funding and skilled labour).

According to Neely (2007), while other PM systems posit customer satisfaction as an antecedent of shareholder value creation (profitability), these do not consider the importance of stakeholder contribution—the contributions the customers and other stakeholders make towards helping the organisation to achieve its goals. Taking this into account Neely, included the fifth dimension Stakeholder Contribution, to emphasise the reciprocal relationship between the organisation and the stakeholders.

¹³ As discussed earlier (section 3.3.2), the BSC starts with strategies (e.g. growth, market share) aimed at creating long-term shareholder value and then identifies key stakeholders to satisfy them to achieve long-term shareholder value.

Although the evidence of practical utility of the PP has been limited (Najmi, Etebari, & Emami, 2012; Smulowitz, 2015), Stakeholder Satisfaction seems to be a valid proxy for social impact, in the case of a NPO.

The PP is not without its critics. Bryson (2011) contends that the PP does not form a link between the strategic indicators and the operational indicators. Bryson also argues that the PP does not provide insights as to how it should be practically implemented for performance improvement. Mendibil and MacBryde (2005) argue that PP is a weak theoretical model because the causal connections between the five performance dimensions have not been well articulated. The latter reason might explain why extant literature on the PP is mostly confined to articles produced by its inventors (Neely and his colleagues). It is important to note that the PP is prescribed as being suitable for an organisation belonging to any sector: private, public, and nonprofit (Neely, 2007).

3.3.3.1. PP Vs BSC: Researcher's Synthesis

The key feature in the PP, according to Neely as well as the proponents of the PP (e.g. Forcada et al., 2008; Najmi et al., 2012), is Stakeholder Satisfaction. The critics of the BSC and the proponents of the PP claim that Stakeholder Satisfaction is not adequately considered in the BSC. The researcher does not fully agree with this criticism on the BSC, as far as for-profit organisations are concerned (the original intent of the BSC was to guide for-profit organisations in strategy monitoring and control). Similarly, the researcher does not view the PP being more aligned to the stakeholder theory than the BSC. Kaplan (2008), the co-creator of the BSC has articulated the conceptual underpinnings of the BSC elegantly (see section 3.3.2). Kaplan argues that the BSC recognises the value of stakeholders for creating long-term shareholder value.

It is clear to the researcher that founders of the BSC (Kaplan and Norton) understood stakeholder theory well, even though they did not embrace it. Encouraging managers to draw strategy maps suggests that the founders of the BSC value the systems thinking concept embodied in stakeholder management. The researcher views the founders of the BSC as *instrumental stakeholder theorists* (see section 2.3.1.2 in Chapter 2) because, they recognised that to maximise long-term value, a firm must manage its stakeholders effectively. What the founders of the BSC argue is that a firm's strategies

should be driven by shareholder needs and not by stakeholder needs. However, in the case of the nonprofit BSC, Kaplan downplays the emphasis on the financial dimension as mission achievement takes precedence over financial performance.

3.3.4. The Conceptual Foundation of the Baldrige Excellence Framework and the EFQM Excellence Model

3.3.4.1. Baldrige Excellence Framework

The Baldrige Excellence Framework (BEF), previously known as “Criteria for Performance Excellence”, and prior to that, “Malcom Baldrige National Quality Award Criteria”, is a performance management framework that serves several purposes: recognition of outstanding organisations (for national or regional quality awards), performance monitoring (e.g. self-assessment), dissemination of best practices, and benchmarking (Flynn & Saladin, 2001; Grigg & Mann, 2008; Saunders et al., 2007). Although the BEF remains a property of the United States,¹⁴ it is being used worldwide for the purpose of gaining benefits that it is supposed to provide (Grigg & Mann, 2008). Although the BEF was originally developed for for-profit organisations, it was extended later (since 2006 on pilot basis and since 2007 permanently) to cover NPOs also, with superficial modifications to its elements; for example, minor changes to the labelling of some performance dimensions (categories) (Jayamaha, Grigg, & Mann, 2009).

The BEF consists of seven performance dimensions (categories): Leadership; Strategy; Customers; Measurement, Analysis, and Knowledge Management; Workforce; Operations; and Results. Each dimension (category) in the BEF is divided into two or more sub-elements known as (measurement) items. The current version (2015-2016) of the BEF contains seventeen items (NIST, 2015). BEF is based on nonprescriptive criteria on performance excellence, in the sense, apart from prescribing how much weight each measurement item should be allocated in calculating the overall performance of the organisation, it only provides broad directions (covered in the next paragraph) on each of the seven performance dimensions (categories) to enable the framework/criteria to be adopted to suit any organisational context (NIST, 2015).

¹⁴ The researcher uses the phrase “property of the United States” to mean that the BF is being developed by the National Institute of Science and Technology (NIST), USA, in conjunction with the American Society of Quality (ASQ), for the benefit of US organisations (mainly to increase their productivity). NIST remains the custodian of the BF.

Figure 3.6 exhibits the current version of the BEF. The basic proposition embedded in the BEF is that Leadership drives the system (People, Information, Strategies and Operational Processes) to achieve results creating a customer focus (Collier & Evans, 2015; Hossain & Prybutok, 2014; Prybutok, Zhang, & Ryan, 2008; Wilson & Collier, 2000).

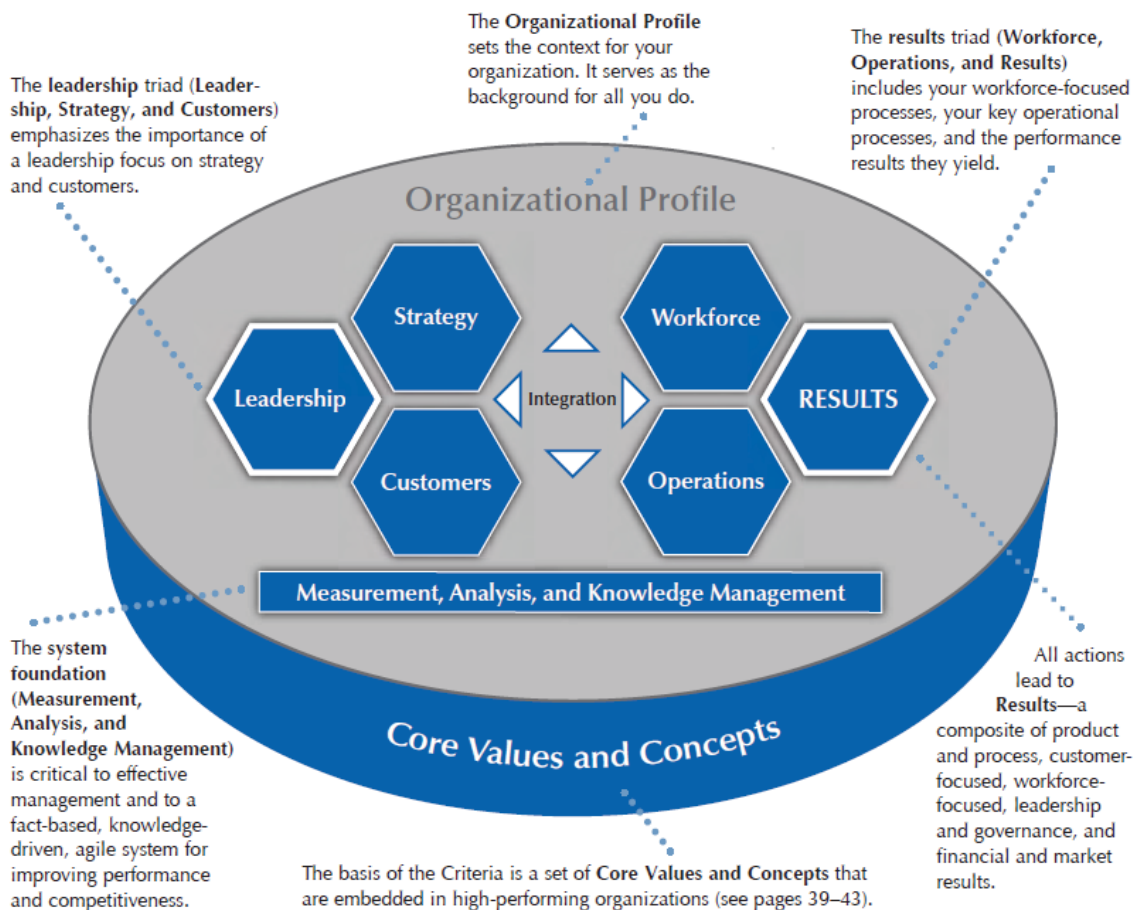


Figure 3.6: The Baldrige excellence framework (NIST, 2015, p. 1)

The BEF can be viewed as a set of nonprescriptive criteria on performance excellence because it raises a series of “what” and “how” type organisation-specific questions¹⁵ on the seven performance dimensions. These seven dimensions are the core of the BEF because they, together, explain how an organisation achieves its higher-level (i.e. strategic) results. According to NIST (2015), because the questions work together as a unique, integrated performance management framework, answering the questions helps

¹⁵ As an example, one of the “how” type questions under the performance domain Leadership is “how do you evaluate the performance of your senior leaders, including the chief executive, and your governance board?” (NIST, 2015, p. 8).

the organisation to align their resources; identify strengths and opportunities for improvement; improve communication, productivity, and effectiveness; and achieve their strategic goals.

The first category of the BEF Leadership examines how senior leaders' actions guide and sustain the organisation, how they communicate with the work force (employees) and encourage high performance; the organisation's governance system and approach to leadership improvement, ethical behaviour and social responsibility are also questioned (examined) (NIST, 2015).

The second category Strategy examines how the organisation establishes its strategy in addressing the challenges (threats and organisational weaknesses) and leveraging on the advantages (opportunities and organisational strengths) (NIST, 2015).

The third category Customers examines how the organisation engages with its customers. This includes how the organisation listens to the voice of its customers, builds customer relationships, and uses customer information to improve and identify opportunities for innovation (NIST, 2015).

The fourth category Measurement, Analysis, and Knowledge Management examines how the organisation selects, gathers, analyses, manages, and improves its data, information, and knowledge assets and how it manages its information technology. This category also examines how the organisation uses feedback to improve its performance (NIST, 2015).

The fifth category Workforce examines the organisation's ability to assess workforce capability and capacity needs and build a workforce environment that is conducive to high performance (in other words human resource planning and development). This category also examines how the organisation engages, manages, and develops its workforce to utilise their full potential to align with the organisation's overall mission, strategy, and action plans (NIST, 2015).

The sixth category Operations examines how the organisation designs, manages, and improves its work systems and work processes to deliver customer value and achieve

organisational success and sustainability. This category also examines employees' readiness for organisational change (NIST, 2015).

The seventh and final category Results examines the organisation's performance and improvement in all key areas: product and process outcomes, customer-focused outcomes, workforce-focused outcomes, leadership and governance outcomes, and financial outcomes. The organisation is required to document its performance levels relative to those of competitors and other organisations with similar service offerings (NIST, 2015).

According to NIST (2015), the BEF can guide for-profit and NPOs to enhance productivity, effectiveness, efficiency through customer satisfaction, workforce satisfaction (volunteers also in the case of a NPO), financial sustainability, governance and social responsibility.

The theoretical validity of the BEF has been examined by number of researchers (e.g. Evans & Mai, 2014; Flynn & Saladin, 2001; Jayamaha et al., 2009; Mellat-Parast, 2015; Pannirselvam, Siferd, & Ruch, 1998). However, there are no large sample studies involving NPOs. Another issue is having a common set of criteria (albeit being flexible enough to be adopted in any organisation) for performance excellence for all types of organisations (Al-Tabbaa, Gadd, & Samuel, 2013; Striteska & Spickova, 2012) implying that BEF does not posit a theory that is unique to NPOs.

3.3.4.2. EFQM Excellence Model

The EFQM Excellence Model (EEM) is the European equivalent of the US's BEF (Bou-Llugar, Escrig-Tena, Roca-Puig, & Beltrán-Martín, 2009; Gómez, Costa, & Martínez-Lorente, 2011; Grigg & Mann, 2008). EEM consists of nine performance domains (Figure 3.7) and each domain is examined much the same way as in the BEF. EEM posits that: (a) Leadership drives the process inputs (People, Strategy, Partnerships and Resources), and (b) Processes deliver outcomes (results) on People, Customers, and the Society and these results together contribute to the Business Results of the organisation (Bou-Llugar et al., 2009; Calvo-Mora, Leal, & Roldán, 2005; Eskildsen & Dahlgaard, 2000). Like the BEF, EEM emphasises the importance of

People and consider it as an “Enabler” of “Results” through process improvement (Paton, 2003); thus unlike the BSC, EEM (and the BEF) recognises the employee contribution for organisational success and sustainability more readily (Al-Tabbaa et al., 2013; Neely, 2007).

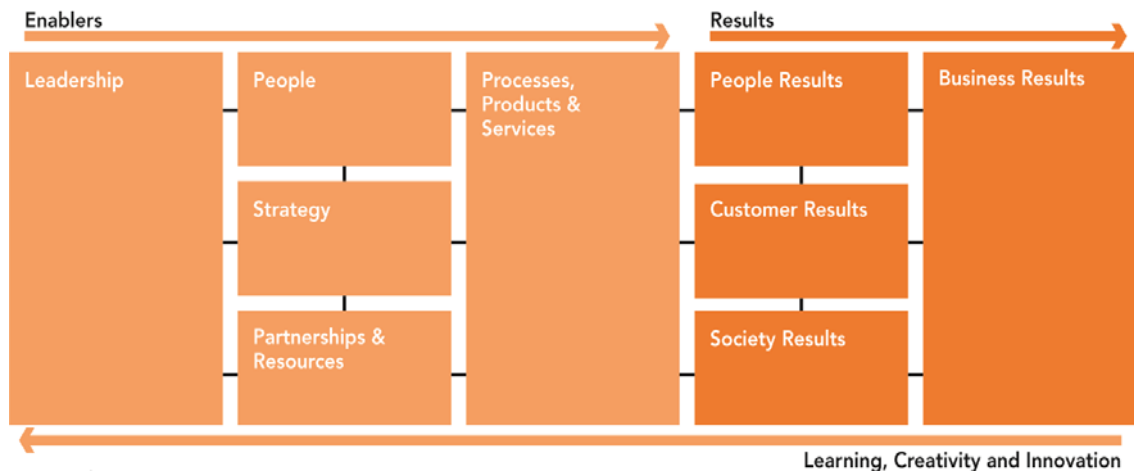


Figure 3.7: The EFQM excellence model (EFQM, 2013)

The theoretical validity of the EEM has been examined by a number of researchers (e.g. (Bou-Llusar et al., 2009; Calvo-Mora et al., 2005; Gómez et al., 2011), although as in the case of BEF, there is no empirical study on the validity of EEM for a sampling frame involving NPOs only. Again, like the BEF, EEM is generic in that the same model is being used for assessment of for-profit and NPOs (Al-Tabbaa et al., 2013; Striteska & Spickova, 2012). For this reason the researcher argues that the EEM does not provide any better theory to explain strategic outcomes on NPOs than the theory provided by the BEF.¹⁶

¹⁶ What these frameworks overlook is discussed in a subsequent chapter (Chapter Eight), after comparing these models with the researcher’s theoretical model.

3.4. PERFORMANCE MEASUREMENT FRAMEWORKS FOR NPOS

The four PM frameworks covered in the previous section (section 3.3) are also relevant to NPOs. These frameworks as well as two other frameworks are reviewed in this section, in relation to NPOs.

3.4.1. The Balanced Scorecard for NPOs

By the time a separate BSC framework was developed by Kaplan (2001) for NPOs, the BSC had progressed towards being a mature framework (the third generation BSC has been published). The nonprofit BSC framework (Figure 3.8) can be viewed as an adjustment to the regular BSC (for-profit version), while retaining the core elements of the third generation BSC, such as the strategy maps. Although the nonprofit BSC retains the four dimensions in the regular BSC, there are differences between the two versions. The major differences between the two versions are that in the nonprofit version: (a) the mission is treated as the driver (as opposed to shareholder objectives as in a for-profit organisation) that sets the strategy, (b) more flexibility being given in forming the dimensions on which the performance measures should be developed (instead of the four rigid measurement dimensions, a nonprofit BSC can have three to five dimensions, depending on the strategy of the NPO being considered), and (c) mission achievement is (as opposed to long-term shareholder value creation) treated as the ultimate goal (endpoint) in nonprofit performance monitoring (Kaplan, 2008).

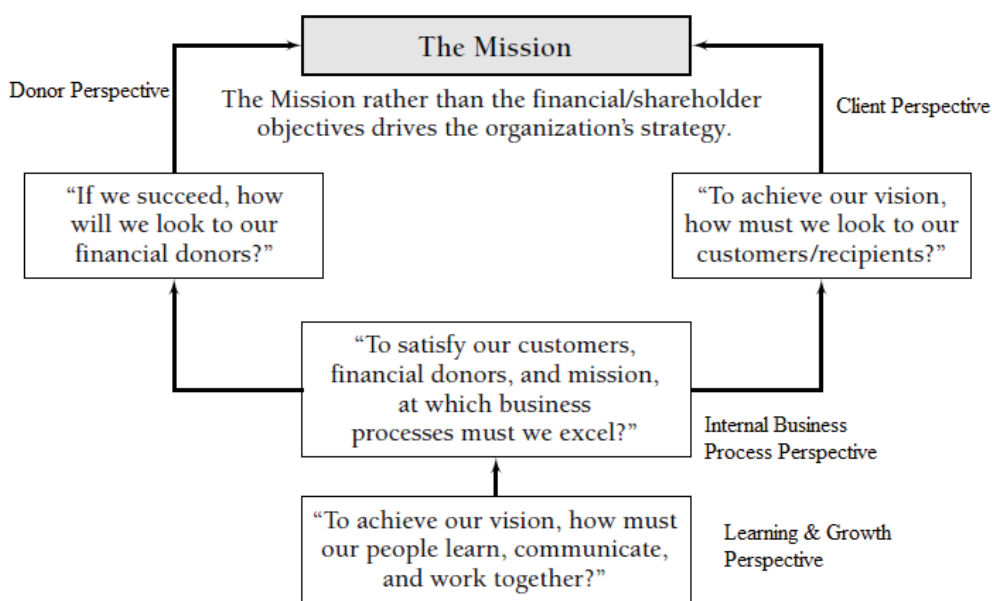


Figure 3.8: The nonprofit BSC framework (Source: Kaplan, 2001, p. 361)

It has also been argued that the customer perspective in the BSC imposes some difficulties in adapting the BSC to NPOs. Unlike the for-profit sector, the nonprofit sector attracts different kinds of customers (Moore, 2003; Talbot, 2010). One customer segment consumes goods and services without paying (beneficiaries) whilst a third party, which is also a customer segment, acts as charitable donors who supply the resources to serve the first customer segment (Moore, 2003; Talbot, 2010). Moore (2003) argues that both customer segments are seemingly equally important to NPOs because fulfilling the aspirations and requirements of both customer segments are crucial for the sustainability of the NPO. Therefore one challenge documented in the literature in implementing the BSC framework to NPOs is the difficulty in identifying who the actual customers are and whether the customer segment *beneficiaries* should be given priority or otherwise over the customer segment *donors* (Anon, 2008; Moore, 2003; Talbot, 2010).

The researcher observes that the nonprofit BSC framework (Figure 3.8) fails to provide a convincing explanation as to why the financial perspective exists and/or the role being played by it (the long-term shareholder value creation concept does not apply for NPOs because they do not have private owners). The role being played by the financial perspective in a nonprofit BSC is also not apparent in strategy maps being prescribed by Kaplan (2008). For example in Figure 3.9, there are no outgoing arrows originating from the financial perspective, which implies that internal operations of a NPO do not depend on its financial perspective! The nonprofit BSC accepts that the financial perspective is necessary for mission achievement, but it fails to explain how achievement of financial objectives causes achievement of the mission and/or the social impact.

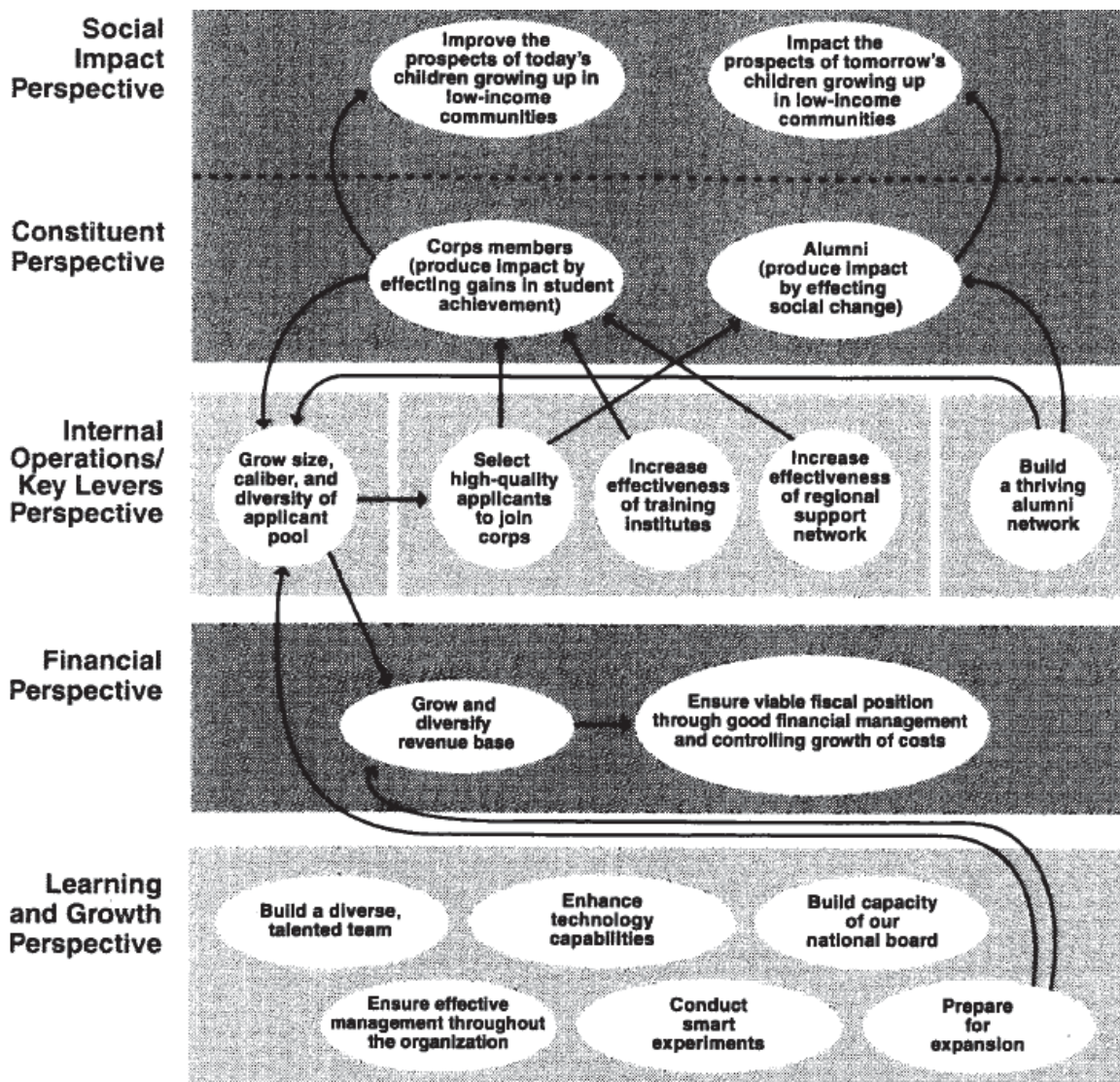


Figure 3.9: A strategy map of a NPO (Source: Kaplan, 2002, p. 233)

Niven (2008) rearranged the BSC perspectives to (among other things) better delineate the role being played by the financial perspective of a nonprofit BSC (Figure 3.10). Niven argued that achievement of employee learning and growth as well as building organisation's physical and knowledge infrastructure (e.g. ICT) is partly dependent on the financial perspective and that in a NPO, the finances can be managed through revenue creation initiatives as well as cost control initiatives (e.g. elimination of waste).

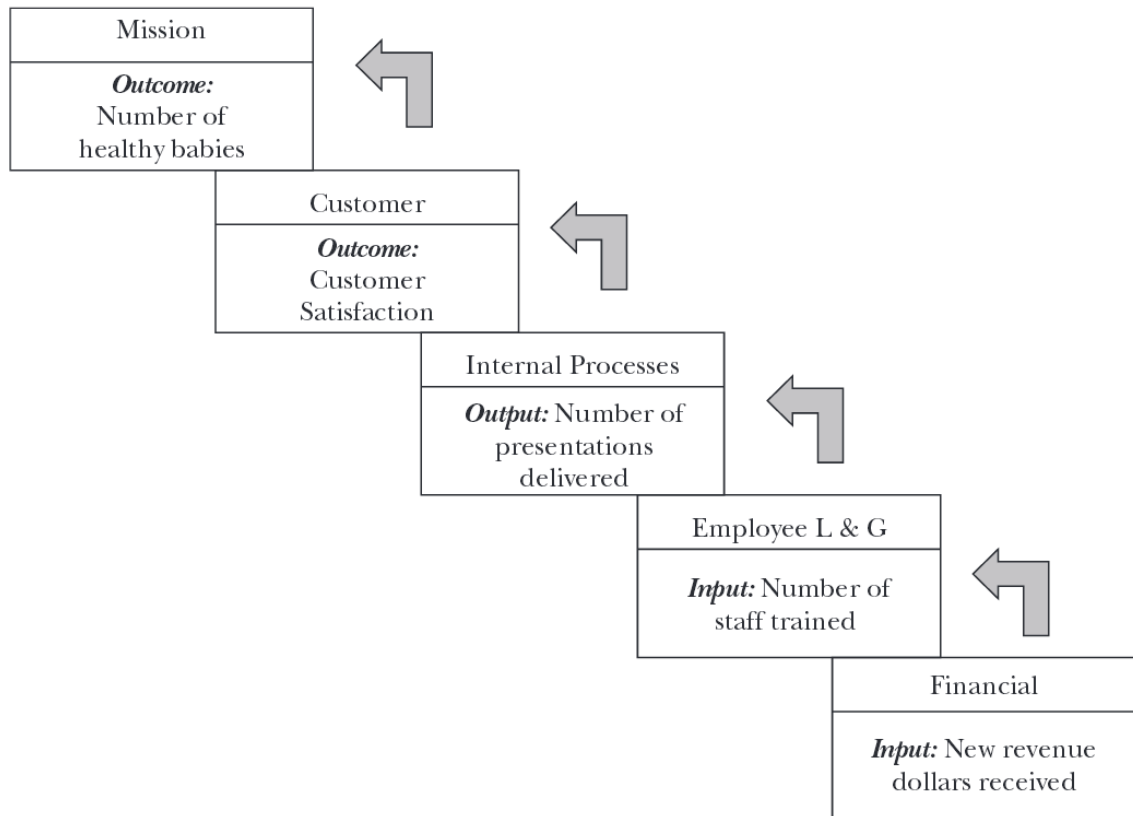


Figure 3.10: The hierarchical arrangement of the BSC perspectives for a NPO as argued by Niven (2008, p.212)

The researcher also observes that the label “Mission” has been used in the nonprofit BSC to mean both the driver (e.g. Trotta, Cardamone, Cavallaro, & Mauro, 2013; Yang, Cheng, & Yang, 2005) and the final outcome (e.g. Kaplan, 2001; Niven, 2008). In the theoretical framework posited by Trotta et al. (2013), the mission is taken as the driver (this is in line with Kaplan, 2001) because mission sets strategies and strategies drive inputs and processes (the BSC perspectives) to achieve the mission-related outcomes. An interesting point in question is if mission is labelled as the driver, what precisely does the Mission represents. Is it the currency and/or relevance of the organisation’s mission? and/or is it the organisation’s commitment to its mission? or is it something else? Although not a research question, the researcher answers these questions indirectly in her fieldwork, in operationalising the concept “Mission”.

3.4.2. The Performance Prism and Quality Award/Excellence Frameworks

The PP (see section 3.3.3 for details), which provides five interrelated dimensions for performance management (Neely et al., 2001) can also be used to measure the organisational performance of NPOs (Anon, 2015). The PP takes a more holistic view of stakeholders than the other frameworks (Tangen, 2004), which is important in a nonprofit context (Neely et al., 2001; Tangen, 2004; Yadav & Sagar, 2013). In keeping with the stakeholder theory, Neely et al. (2001) emphasise the importance of taking stakeholder requirements into consideration before establishing the organisational strategy. Tangen (2004) argues that questioning the organisation's existing strategy in relation to the stakeholder environment provides a strong foundation in selecting more relevant performance measures. Yadav and Sagar (2013) argue that a stakeholder focus on strategy formulation (rather than a shareholder focus) creates a long-term focus on organisational value creation.

While the above arguments (i.e. Neely et al., 2001; Tangen, 2004; Yadav & Sagar, 2013) have not necessarily been directed in support of NPOs, it seems to the researcher that the stakeholder focus advocated in the PP seems to better explain why people are motivated to work in NPOs (many work voluntarily for non-monetary rewards in NPOs and unless a NPO creates an employee focus, it will fail to retain its workforce) and why NPOs exist in the first place. This is because theory of the firm (McWilliams & Siegel, 2001), which forms the basis of the BSC's shareholder view, does not fit NPOs (there is no such thing as maximising the net benefits of the future returns of the shareholders, in the case of NPOs).

Al-Tabbaa et al. (2013) opine that both quality award frameworks such as the BEF and EEM (these were covered earlier in section 3.3.4) are generally applicable to NPOs because they take a stakeholder perspective. For example, they show that the measurement domain Society Results in the EEM (see Figure 3.7) captures the importance of NPO's outcomes on the society, which is important in attracting donors, which is turn vital for long-term sustainability (Al-Tabbaa et al., 2013).

3.4.3. The Production of Welfare Framework

Production of Welfare Framework (Figure 3.11) is a performance measurement framework that is designed to examine the production of welfare (final outcome) in relation to resources (production inputs) using an analogy similar to that being used in microeconomics to explain the production of conventional goods using production factors (labour and capital) (Davies & Knapp, 1994; Kendall & Knapp, 2000). This framework is one framework that has been tailor made to measure the performance of voluntary organisations, particularly healthcare service organisations (Davies & Knapp, 1994; Iparraguirre & Ma, 2015).

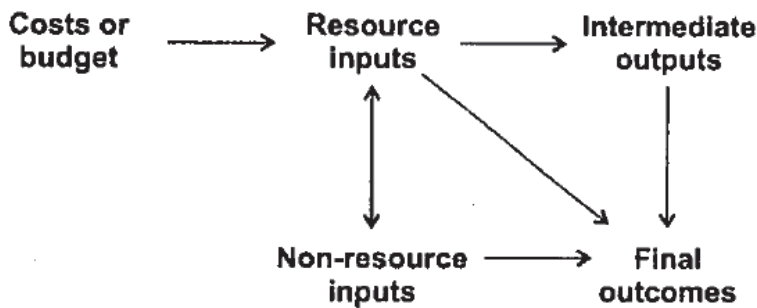


Figure 3.11: Production of welfare framework (Source: Kendall & Knapp, 2000, p. 115)

Five key features are embraced in the Production of Welfare Framework: Resource Inputs (employees, capital); Cost or Budget; Non-resource Input (ideas or attitudes of internal participants); Intermediate Outputs or Outcomes (services provided and their quality); and the Final Outcomes (improvement in the beneficiaries within the assessment period) (Kendall & Knapp, 2000). The *meso* and *macro* context are also entailed in the framework, which embraces the internal and external environment respectively. The framework is further developed using eight performance domains and indicators: Economy, Effectiveness, Efficiency, Pluralism, Equity, Participation, Advocacy and Innovation (Kendall & Knapp, 2000). Of all the PM models, the Production and Welfare Framework seems to be the most difficult model to adopt as there is virtually no explanation as to how the model could be used in strategy monitoring and control. This is partly because the Production of Welfare Framework does not examine the processes that are involved in transforming inputs into outputs; this is not a surprise, given that this framework is grounded in economics.

3.4.4. The Logical Framework Approach/Logic Model

The Logical Framework Approach (LFA) is a management approach that is used for project planning to measure the social impact of a particular project (Kellogg Foundation, 2006). As such the LFA/Logic Model has relevance to any organisation (private, public or nonprofit) as responsible citizenship is common to any project (e.g. planning a coal power station, commissioning a nourishment programme for under-privileged school children) (Hatch & Cunliffe, 2013; Maas & Liket, 2011). Figure 3.12 depicts the Logic Model in its basic form.

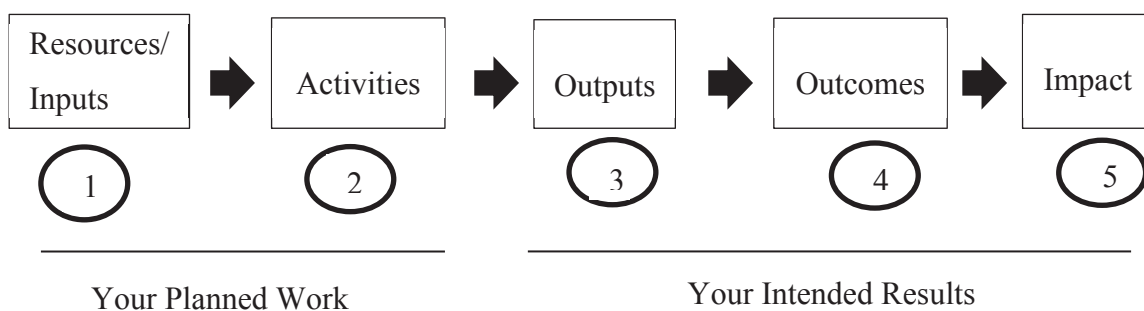


Figure 3.12: The basic Logic Model (Kellogg Foundation, 2006)

The fact that the Logic Model addresses the social impact of a project makes it a viable model for PM in NPOs. The downside of the Logic Model is that the model measures performance at “project level” as opposed to “organisational level” (Maas & Liket, 2011). The model reminds planners that they need to look beyond the system as what eventually matters is the (social) impact the project makes (in a positive or a negative way) on the stakeholders, and not the outputs per se (Maas & Liket, 2011; Mulgan, 2010). Maas and Liket (2011, p. 175) define *Impact* of a project as the changes the project makes to *social systems* minus what would have happened in the absence of the project.

The LFA (the logic model and the associated tools) consists of nine key features: Context, Stakeholder Analysis, Problem Analysis, Objective Analysis, Activities, Resource Planning, Goal Achievement Indicators Identification, Risk Analysis and Assumptions (Bakewell & Garbutt, 2005). The LFA does not prescribe a priority order in which the nine features should be arranged; flexibility is given to contingency factors

such as the organisational requirements. This may mean that the use of the approach may differ from one organisation to another. The logical and important element in the LFA is the 4x4 matrix (the rows in the matrix refer to level of objectives and the columns refer to achievements of the objectives), which gives the summary of the programme's key elements and relationships among them—as well as the overall procedure that allocates program elements (Bakewell & Garbutt, 2005).

In spite of having a workable definition on *social impact* and looking beyond the system to measure social impact, LFA seems to be too simple, lacking essential PM dimensions that are considered to be important in managing organisational performance (Bakewell & Garbutt, 2005). Moreover, being too simple, LFA leads different organisations to use the approach in different ways, thus limiting its generalisability (Bakewell & Garbutt, 2005; Harley, 2005). Therefore LFA is typically used for project planning and evaluation instead of measuring the overall organisational performance (Harley, 2005). In addition, Harley (2005) argues that difficulty in selecting program indicators results in implementation problems of the Logic Model, even at the project level.

3.5. PRIOR STUDIES THAT EXAMINED DIFFERENT FACETS OF VALIDITY OF THE BSC

This section solely reviews prior studies on different facets of validity of the BSC in a *nonprofit context*. With the exception of quality award frameworks, other PM frameworks have received very little academic scrutiny on theoretical and practical validity. This section includes papers on BSC implementation studies (case studies), conceptual papers (usually literature synthesis and/or authors' personal reflection), and large sample studies (these are quantitative studies used for theory testing).

3.5.1. BSC Implementation Studies

The applicability of the BSC to NPOs has been tested at organisational level. For example, using the multi-criteria decision making approach Grigoroudis, Orfanoudaki, and Zopounidis (2012) demonstrated the development of a BSC (including strategy map development) in a healthcare organisation. The most essential financial performance indicators as well as non-financial indicators were embedded in the proposed model and these were used to assess service quality, stakeholder satisfaction, and organisational

improvements. Grigoroudis et al. found that the performance of their case study organisation was low in the Customer perspective, implying that the case study organisation has not met customer expectations due to ineffective service processes as implied by the low performance in the Internal Business Processes perspective.

Martello, Watson, and Fischer (2011) examined how the implementation of the BSC enables bridging the gap between operational activities and strategic objectives in a rehabilitation centre in New York, USA. Based on information collected from the respondents, they observed that equal importance needs to be given to the Customer dimension as well as the Financial dimension, to achieve the Mission.

Through conducting interviews with a select group of personnel (the vice president, the administrative team and the board of directors) Kocakulâh and Austill (2007) studied the Midwestern healthcare system in the US to answer the research question “why should the BSC be used by healthcare organisations?” They found that the four PM perspectives prescribed in the BSC do not adequately fit the nonprofit healthcare sector in explaining healthcare performance and hence the BSC perspectives need to be substantially redefined as Customer Satisfaction, Clinical Quality, Operations, and Financial to suit the sector; the researchers also emphasised the importance of modifying the BSC according to the mission and vision of the organisation before it is being adopted for strategic management purposes (planning, implementing, and monitoring the strategies).

By proposing an implementation framework of the BSC, Yang et al. (2005) attempted to validate the BSC through a case study involving a Taiwanese nonprofit hospital. Their model consisted of five PM dimensions: Customer, Internal Business Process, Learning and Growth, Financial, and Society Promise. They concluded that the BSC fits NPOs but they emphasised the importance of continuing research in other NPOs for further refinement.

Stewart and Bestor (2000) conducted a case study on a nonprofit Community Memorial Hospital in Washington and Ozaukee counties in Wisconsin, USA to design and implement an integrated PM system for their case study hospital as this hospital was facing both market deregulation and competition for customers (patients) from other

hospitals. Stewart and Bestor employed the BSC to develop an integrated PM system, which contained several elements: performance criteria, standards and measures and the link between hospital objectives, operating plans and day-to-day performance. They found that the BSC approach to PM provides an effective and efficient way to combine day-to-day activities with the strategic objectives of the organisation.

While there have been no Australasian studies on the relevance or validity of the BSC in NPOs, there have been interesting studies (e.g. Greatbanks & Tapp, 2007; Hoque & Adams, 2011; Northcott & Taulapapa, 2012; Ridwan, Harun, & Fahmid, 2013) on the applicability of the BSC in the public sector organisations such as city councils, secondary and tertiary hospitals belonging to the DHBs, government departments and (Australian) state owned enterprises. The gaps found in some of the studies have relevance to NPOs as the public sector organisations that were covered in these studies operate in a somewhat similar fashion to some nonprofit Australasian organisations (e.g. the social impact is important to both sectors).

Greatbanks and Tapp (2007) conducted a longitudinal case study on the Dunedin city council to determine how the BSC impacts on the council's employees' activities and processes to deliver customer service. They found that the BSC has a positive impact on the employees in executing the council's strategic plan to deliver customer service.

Hoque and Adams (2011) studied the breadth and depth of application of the BSC in Australian government departments. They found that Australian government departments that use the BSC as a primary mechanism to implement the strategy outperform the departments that do not use the BSC intensely for strategy implementation. What is not clear in the study is whether the departments that do not use the BSC intensely do use any other formal PM system. Otherwise one can argue that the performance effect has come from formal performance monitoring rather than the BSC.

Using a postal survey¹⁷ followed up by an interview of a select group of personnel Northcott and Taulapapa (2012) studied how the BSC is being implemented in New

¹⁷ The authors have used the postal survey and the follow-up interview to gather qualitative data (hence this study is not classified as an empirical study).

Zealand government departments to identify shortcomings of the BSC as a tool for performance management in the public sector. They found that the four performance perspectives prescribed in the BSC need to be substantially redefined to suit the sector and that causal links between the perspectives need to be better articulated for managers to use the BSC with greater conviction.

3.5.2. Conceptual Papers

Kong (2010) argued that as a strategic management tool, the BSC is less effective in delivering the social value a NPO is expected to deliver, due to incompatibility of the four performance perspectives with the strategic goals of a NPO. In particular he argues the nonprofit BSC does not address certain key aspects in a nonprofit strategy, such as workforce elements (e.g. attracting and retaining the workforce), social elements, and factors that are unique to the external environment of a NPO (e.g. specific types of collaboration and completion). He argues that one approach to circumvent these shortcomings of the BSC is to introduce the “Intellectual Capital Framework” (Youndt, Subramaniam, & Snell, 2004) in NPOs to supplement the BSC.

Through literature review, Bisbe and Barrubés (2012) examined the BSC as a strategy implementation and monitoring tool in healthcare organisations. They identified three areas of concern that need further research: (a) the need to operationally describe the PM dimensions to suit healthcare, (b) the difficulties of implementing the BSC due to having to have *strategy maps* as a prerequisite in the third generation BSCs and (c) the causal order in which the PM dimensions need to be arranged.

Through literature review, Trotta et al. (2013) proposed a BSC framework to assess the strategic performance of teaching hospitals. Their conceptual framework posits that the Mission and Vision drive strategic goals, which in turn drives the Innovation and Growth perspective, the Research perspective, the Teaching perspective, the Patient Care perspective, Economic/Financial perspective, and the Stakeholder perspective. The researcher observes that the Research perspective and the Teaching perspective (sub-processes that are relevant to teaching hospitals) are not relevant to healthcare NPOs in her study.

The UK research organisation 2GC Limited identified five factors that hinder development of scorecards in the nonprofit sector: difficulty in operationalising what *success* means, difficulty in changing the scorecard frequently as the organisation adjusts and readjusts to strategic challenges, difficulty in defining measurable performance measures, the need to change the scorecard to take into account “limited accountability for results”, and lack of guidance on how financial metrics should be developed (Anon, 2008).

3.5.3. Empirical Studies

Using literature on the BSC to formulate a theoretical model and structural equation modelling (SEM) to test the model, Wu and Hung (2008) conducted an empirical study ($N = 135$) to test the effect of “cause-related marketing”¹⁸ on the mission achievement of a single Taiwanese NPO. They examined the BSC performance factors (constructs) and the hypothesised theoretical relations between the factors based on data collected from the NPO. They hypothesised that Learning and Growth causes Internal Processes which in turn simultaneously causes Financial Sustainability and Customer Satisfaction—both of which in turn cause mission achievement. Out of the five constructs they modelled, two constructs—Learning and Growth and Internal Processes—were represented as multidimensional (second-order) constructs that reflect several first order constructs; Learning and Growth was represented via Human Capital, Organisational Capital and Information Capital (in essence resources) while Internal Processes was represented via Operations Management, Regulatory and Social Management, Customer Management, and Innovation. The researcher observes that using this number of constructs¹⁹, which also means so many measures to represent a BSC measurement perspective such as Internal Processes, is not consistent with the principles of the BSC—particularly, the bounded rationality principle. Increased level of abstraction (second order constructs are very abstract) is also a practical issue

¹⁸ Cause-related marketing (CRM) is a distinct marketing strategy that can be adopted by a for-profit organisation to achieve for-profit objective through charitable giving. The following formal definition of CRM was proposed by Varadarajan and Menon (1988, p 60): “*Cause-related marketing is the process of formulating and implementing marketing activities that are characterized by an offer from the firm to contribute a specified amount to a designated cause when customers engage in revenue-providing exchanges that satisfy organizational and individual objectives*”.

¹⁹ Though not explicitly stated by the authors, part of the reason for using large number of constructs could be to capture certain initiatives on the part of the NPO that “cause-related marketing” would have supposed to address.

because to the practitioner, any framework that academia use must be comprehensible. Moreover, the study involved only one organisation (multiple respondents), meaning low generalisability. Wu and Hung concluded that a better theoretical model needs to be explored to study the effect of cause-related marketing on a NPO.

Valderrama, Cornejo, and Bordoy (2013) used the data envelopment analysis method to validate the BSC in a Spanish university. They collected empirical evidence from academic management units to validate the implied causal relationships between the BSC's constructs pertaining to the university, and to design the strategy map for their case study organisation. They argued that their proposed BSC model is suitable for use by the departments of the University to measure their internal performance and the relationship of these to the resources that they (the departments) have used.

Adopting a multiple case design to collect qualitative data (observation of board meetings, interviews with directors and CEOs, and access to board documents), Hough, McGregor-Lowndes, and Ryan (2015) examined the scorecard's impact on board monitoring practices in five small and medium-sized healthcare NPOs in Queensland, Australia to answer the research question: "how does the availability of a BSC-type tool impact on board monitoring in nonprofit organisations?" From their findings (passive usage of the BSC was one of their main findings), they concluded that there's no evidence to support the proposition that a BSC-type tool results in improved understanding of organisational performance by the board, although Kaplan and Norton (1996b, 2001) claimed so. Moreover, they found that there are several challenges in adopting the BSC (in its default form) as a board-monitoring tool. They concluded that if BSC is to be an effective board monitoring tool, a NPO must establish appropriate data collection and data analysis routines and practice these routines diligently, as part of the board's culture.

3.6. KNOWLEDGE GAPS AND RESEARCH QUESTIONS

In the previous chapter (Chapter Two), among other things, the researcher examined the nature of the nonprofit sector (through a literature review) to argue that there are differences between managing a NPO and a business (for-profit organisation), or even a public sector organisation. This shows that a PM system that addresses the requirements of a for-profit organisation is probably unlikely to adequately address the requirements of a NPO, and therefore, use of such systems could result in ineffective or suboptimal outcomes for NPOs.

This chapter examined the literature on key PM frameworks to examine how they stand, in relation to NPOs. The literature review revealed that nonprofit BSC provides some key elements that are important to nonprofit management (see section 3.4 for details). Firstly, it recognises that Mission of the NPO is what drives the organisation's strategy, and not primary shareholder objectives such as growth and profitability (Kaplan, 2001, 2008). The BSC also takes a systems perspective in that it recognises that it is the strategic resources and organisational capabilities (Learning and Growth perspective) that enable the Internal Process to deliver NPO's mission-related outcomes. By induction, it follows that it is the strategy of the NPO that causes the development of its strategic resources and organisational capabilities; principles of strategic management (e.g. Eden & Ackermann, 2013; Grant, 1991) also supports this notion.

Niven (2008) provided more clarity to the nonprofit BSC (see section 3.4.1). In particular, he showed that financial performance (Financial perspective in the BSC) is not a mission-related outcome as envisaged by Kaplan (2001), but a requirement that is needed for developing the intangible resources (Learning and Growth perspective in the BSC). He also argued that in a NPO, financial performance can be improved by taking strategic initiatives on revenue gathering and cost control. In Chapter Two the researcher showed that NPOs gather revenue from philanthropy and the government, and in the case of some NPOs, also charging fees from the client (Salamon et al., 2013).

In section 3.5, the researcher examined relevant prior studies on the nonprofit BSC, particularly in relation to healthcare. The general finding was that the four measurement perspectives prescribed in the BSC (plus Mission, in cases where it has

been considered such as in the study by Trotta et al. (2013)) do not adequately represent nonprofit management (e.g. Bisbe & Barrubés, 2012; Kocakülâh & Austill, 2007; Kong, 2010). Lack of full consideration of key stakeholders (e.g. Neely et al., 2003; Yadav & Sagar, 2013), and in particular, overlooking the workforce as a key stakeholder group (hence non-allocation of a dedicated measurement perspective on people-related outcomes) was one of the major shortcomings of the BSC highlighted by the scholars (e.g. Kong, 2010; Pessanha & Prochnik, 2006). Another criticism levelled at the BSC (e.g. Bisbe & Barrubés, 2012; Nørreklit et al., 2012) was found to be lack of clarity on causal relationships between BSC's performance dimensions (perspectives).

At a practical level, having to draw strategy maps to rationalise the selection of performance measures was found to be a deterrent in adopting the BSC by managers (e.g. Bisbe & Barrubés, 2012; Dechow, 2012) as managers need significant input from external consultants (Hoque, 2014; Qu & Cooper, 2011; Noell & Lund, 2002). Another criticism of the BSC, which has relevance to both theory and practice is lack of clarity on how performance measures should be developed (e.g. Parmenter, 2012) and limiting the number of constructs (performance perspectives or dimensions) to four (and mission when it is operationalised) and imposing limitations on the number of performance measures that the BSC should accommodate (e.g. Salterio, 2012).

The PP (section 3.3.3) and Quality Award/Excellence Frameworks such as the BEF and EEM (section 3.3.4) addressed some issues of the BSC. For example, these frameworks take a more holistic view on the stakeholders. In addition, the BEF and EEM provide an overall performance scoring system. However, the PP and Quality Award/Excellence Frameworks appeared to be generic prescriptions for performance monitoring and improvement for any organisation in any sector (for-profit, public, and nonprofit), which defeats the researcher's fundamental argument that nonprofit management is different from for-profit (and even public) management.

Having identified the knowledge gaps pertaining to the nonprofit BSC, and having identified the properties of other PM frameworks, the challenge is to augment the BSC to better suit NPOs, keeping its key characteristics intact. With this in the backdrop, the researcher proposes the following three research questions.

RQ1. What set of PM dimensions constitute an integrated PM system for NPOs?

RQ2. What are the operational definitions of these PM dimensions?

RQ3. How do the PM dimensions theoretically relate to one another in explaining the achievement of strategic objectives of a NPO?

The researcher also highlighted that the BSC fails to provide an overall indication as to how the organisation is performing. Stated alternatively, the BSC fails to explain how the measures as a whole indicate how an organisation is progressing overall, in implementing its strategies (Meyer, 2002). This is because the performance measures in the BSC are unweighted. Kaplan (2008) himself has admitted that this shortcoming may have implications on the stakeholders (e.g. in a for-profit organisation, the shareholders). For a NPO this means that key donors (especially the government) are unable to easily gauge how the leadership (the board of directors) is performing to link funding to performance.

A manager can easily be misguided by random variability of performance over time, based on a single indicator. This is because an organisation as a system, is subjected to numerous randomly varying causes (common causes) that are beyond the control of the organisation (Britz, Emerling, Hare, Hoerl, & Shade, 1997; Deming, 1994; Hoerl & Snee, 2012). Quality Guru W. Edwards Deming spent much of his professional life in educating managers when to take action (when the variability can be attributable to a special cause) and when not to take action (when the variability is random and cannot be attributable to a special cause). In addition, Deming (1994) also showed that random variability (hence common cause variability) does accrue due to imperfections of the measurement system (no measurement system is perfect)—sometimes a measurement being above the true value and other times below the true value. Thus consolidating (aggregating) each performance measure in a PM system *on a rational basis*, to constitute the overall performance of the organisation will filter common causes (noise) from special causes to represent a very useful measure of overall performance. This is not to undermine the fact that managers need to be able to distinguish random variation (as indicated by an individual performance measure) from special cause variation through statistical literacy (Britz et al., 1997; Hoerl & Snee, 2012).

Thus the fourth (final) research question RQ4 is stated as follows.

RQ4: What weighting should managers give to each PM dimension and its subparts in assessing the overall strategic performance of a NPO?

3.7. CHAPTER CONCLUSION

This chapter reviewed the literature on several PM frameworks. The conceptual foundations of the BSC and other PM frameworks were deeply examined through the synthesis of the literature (sections 3.3 and 3.4). In addition prior research on examining the validity of the BSC in a nonprofit context (especially healthcare) was discussed (section 3.5). The literature review resulted in the identification of several knowledge gaps in the nonprofit BSC (concisely summarised in section 3.6). The researcher argued that although PM frameworks such as the PP, BEF, and EEM possess useful characteristics, the nonprofit BSC provides a useful initial theoretical perspective to forge ahead in developing it further to suit NPOs. The challenge is to augment the BSC to better suit NPOs, keeping its key characteristics intact. With this as the backdrop, the researcher proposed her questions (four altogether, as shown in section 3.6). The next chapter (Chapter Four) covers the methodology that was adopted to answer the research questions, and thereby achieve the research objectives.

CHAPTER FOUR

THE RESEARCH METHODOLOGY

What a man sees depends both upon what he looks at and also upon what his previous visual-conceptual experience has taught him to see.

Thomas S. Kuhn

Things are as they are because they were as they were.

Thomas Gold

4.1. INTRODUCTION

The methodology adopted by a researcher to achieve research objectives heavily depends on the beliefs of the researcher, as to what constitutes knowledge (Dey, 2002; Jehn & Jonsen, 2010). Thus this chapter begins (section 4.2) with a description of key research paradigms (knowledge claims) prevailing in social science research. This is followed by a description of mixed methods designs (section 4.3), an approach that goes hand in hand with the research paradigm ‘pragmatism’. In addition, the selection of the research design and justification of researcher’s paradigm (pragmatism), are also covered in this section. The conceptual model (Figure 4.4), that follows from the key literature covered in the previous chapter, is described next (section 4.4). The conceptual model development is followed by qualitative data collection via a case study, for operationalisation of constructs and refinement of the conceptual model (section 4.5). Quantitative data collection to test the models is covered next (section 4.6), followed by a description of methods and techniques that were adopted to analyse the data (section 4.7). This is followed by an explanation of how an ancillary model was developed to determine the empirical weights of each construct, in determining the overall performance of an organisation (section 4.8). This chapter also justified the quality of the research (section 4.9), using quality criteria prescribed in mixed methods research. Finally, the chapter ends with a summary of the methodology that was used to answer the research questions.

4.2. RESEARCH PARADIGMS

Comprehensive, complete, and methodological sophistication is needed for conducting PhD research (Houghton, Hunter, & Meskell, 2012). To this end, it is also important to look through one's philosophical worldview assumptions, the strategy that relates to worldview and the techniques or procedures that one uses to accomplish their research objectives—collectively known as the *research paradigm*. Guba and Lincoln (1994) define a paradigm as a set of beliefs that symbolise researchers' "worldview", their place in it, and connection to that world. A paradigm consists of three components:

- Ontology – basic beliefs about the nature of reality.
- Epistemology – nature of the relationship between researcher and what can be known.
- Methodology – the process of how the researcher seeks out new knowledge, the main beliefs of research and how it should proceed (Denzin & Lincoln, 2011).

There are a plethora of sources that describe research paradigms and their components (Creswell, 2014; Denzin & Lincoln, 2011; Guba & Lincoln, 1994; Johnson & Onwuegbuzie, 2004). It is important to have a working knowledge of relevant paradigms to justify the research design (Houghton et al., 2012). Denzin and Lincoln (2011) describe three major paradigms being used in research: *positivism/postpositivism*, *constructivism (interpretivism)*, and *participatory (advocacy)*. Apart from these three paradigms, there is another important emerging paradigm known as *pragmatism*, which results in so-called "mixed methods research" (Creswell, 2014; Easterby-Smith, Thorpe, & Jackson, 2012; Johnson & Onwuegbuzie, 2004). These paradigms are briefly described in turn. Justification of the paradigm of the researcher (pragmatism) is given later (section 4.3.3).

4.2.1. Positivism and Postpositivism

The positivistic ontology holds that "there is a reality out there to be studied, captured and understood" (Guba, 1990, p.22). The positivist ontology also holds that the researcher (observer) and the subject (e.g. the observed person) are independent (detached). The positivistic epistemology holds that the role of the researcher is to examine this reality—which exists out there irrespective of the observer—by testing hypotheses (mostly cause-effect in nature), by way of making precise measurements

related to the phenomenon under observation (Creswell, 2014; Denzin & Lincoln, 2011; Easterby-Smith et al., 2012). Much of knowledge of the sciences (not necessarily the case in social sciences) has advanced by testing hypotheses. This way of advancing knowledge is known as “hypothetic-deductive” reasoning in the academic literature (Orlikowski & Baroudi, 2002). Both confirmation of hypotheses and falsifying hypotheses enable a scientist to advance knowledge (Orlikowski & Baroudi, 2002; Popper, 2002). For these reasons positivism involves *quantitative data* and associated research methods such as statistical hypothesis testing (Bryman, 2012; Creswell, 2014).

Postpositivism is an extension of positivism. The philosophy behind postpositivism is that outcomes or effects influenced by causes can only be judged (imperfectly) but never be fully captured (Guba, 1990); post positivism also embraces hypotheses testing but it emphasises the importance of finding the context of the cause through “triangulation”, using evidence from multiple streams of data (Creswell, 2014; Guba, 1990). Myers (2013) observes that positivism (postpositivism included) remains the dominant research paradigm in many businesses and management research disciplines, although other paradigms have received increased attention more recently.

The rigour of positivist and postpositivist research can be assessed through four quality criteria: reliability, construct validity, replicability, and external validity (generalisability) (Creswell, 2014; Wahyuni, 2012). The consistency of the measures that are used to define the constructs is defined as reliability (Cronbach, 1951; Nunnally & Bernstein, 1994; Wahyuni, 2012) while the extent to which a measurement scale measures what it is required to measure is defined as construct validity (Nunnally & Bernstein, 1994). The ability of a research to be repeated and obtain similar outcomes can be defined as replicability, while the extent to which the results can be generalised via the targeted population is defined as generalisability (Crook & Garrat, 2011).²⁰

4.2.2. Constructivism

The assumption behind constructivism, also known as interpretivism, is that subjects’ meanings are developed by the researcher based on information and knowledge that have been gained from experience—in other words, constructivism/interpretivism

²⁰ These concepts are discussed in detail later (Chapter Five).

assumes that knowledge is socially constructed (Creswell, 2014; Guba & Lincoln, 1994). From an ontological and epistemological stance, constructivism is the polar opposite to positivism (and postpositivism). Constructivists are guided by the meanings that look for more complex views (deep rich meaning) rather than reducing the meaning down to a set of testable hypotheses. Using the meanings as a base, constructivists establish theories or knowledge claims instead of starting with a theory for testing (Creswell, 2014). Consequently, the concepts that underpin the theories emerge as the researcher interacts with the subjects and hence, “inductive reasoning” is used to advance knowledge in constructivism (Guba & Lincoln, 1994). From a methodological standpoint, the data used for this reasoning is mostly *qualitative*, such as the data collected through observations or interviews engaging in a case study (real-life) setting (Stake, 1995; Yin, 2014). Yin opines that case studies are best suited to when a researcher wants to get a deeper understanding of the area being studied.

It is important to distinguish a case study (a specific method, which can be distinguished from other methods such as surveys and experiments) from a paradigm. Case studies can be conducted within any research paradigm (Darke, Shanks, & Broadbent, 1998; Myers, 2013). In a constructivist case study, a researcher attempts to understand a social phenomenon via the meanings people (informants) assign to it (Myers, 2013; Walsham, 1995). The quality of a constructivist case study is judged by the credibility (trustworthiness) of the story (on the social phenomenon) unfolded by the researcher (Myers, 2013). In a positivistic case study, the focus is on developing testable propositions (hypotheses) to predict and explain a social phenomenon, and positivistic notions of quality criteria apply: construct validity, internal validity, external validity and reliability (Myers, 2013).

Guba and Lincoln (1989) attempted to define parallel notions, which can be considered as generic quality criteria for qualitative validity: credibility, transferability, confirmability, and dependability. Credibility (approximately analogous to internal validity) refers to establishing confidence in the reader on the trustworthiness of the findings (e.g. peer debriefing and member checking) while the transferability (approximately analogous to external validity) refers establishing judgments of the degree to which findings are similar to other contexts. Guba and Lincoln assert that a well-described case study requires a clear description of the context and statement of the

assumptions used by the researcher, so that other researchers who may want to project the findings of the case study to another setting, could use his/her own judgement about the generalisability of the case study findings. Dependability (somewhat analogous to reliability) refers to the extent to which the findings of the study could be relied upon as being stable (Bitsch, 2005), while confirmability refers to the extent to which findings and data can be traced to the sources or the responses but not to the researcher's personal beliefs (biasedness) (Guba & Lincoln, 1989).

4.2.3. Participatory/Advocacy Worldview

Participatory/advocacy worldview holds that marginalised peoples and groups in a society (e.g. women and minority in certain societies) require an “action agenda” that would emancipate them from marginalisation (i.e. social justice) (Bent-Goodley, 2001; Creswell, 2014). Creswell contends that participatory/advocacy researchers follow a highly participatory approach in engaging with their informants (e.g. a marginalised or disadvantaged group in a society), to ensure that their informants do not get further alienated. In justifying their stance, participatory/advocacy proponents argue that although the constructivism worldview is supportive of a participatory approach, it does not adequately fit agendas on marginalised or “disenfranchised” peoples in a society (Creswell, 2014; Green, McCollum, & Hays, 2008). The advocacy paradigm is not relevant to this researcher's research questions and research objectives.

4.2.4. Pragmatism

Pragmatism is a framework for realising the usefulness of “application”, which depends on the area of study and the researcher (Rorty, 1991). Pragmatism involves applications or research questions and answers instead of specific paradigm-locked methods (Creswell, 2014; Johnson & Onwuegbuzie, 2004; Tashakkori & Teddlie, 2010; Wahyuni, 2012). One of the key tenets of pragmatism is that *philosophical activity* should focus on answering the problems and not forming paradigms (Tashakkori & Teddlie, 2010). Pragmatism (the paradigm) goes hand in hand with research designs that involve a mix of qualitative and quantitative data (Creswell, 2014).

4.3. MIXED METHODS DESIGNS

Mixing or combining quantitative and qualitative research practices, methods, approaches, or ideas in a single study or series of studies is defined as “Mixed Methods Research” (Creswell, 2014; Creswell & Clark, 2011; Johnson & Onwuegbuzie, 2004). The mixed methods research espoused in pragmatism has the following positive features:

- It presents a *middle ground philosophy*, which takes the best of both worlds: positivistic (postpositivistic included) and constructivist (Creswell, 2014). Mixed methods research views quantitative data and qualitative data as being complementary (Creswell, 2014; Jick, 1979).
- It provides a realistic and results-oriented method (Creswell, 2014; Johnson & Onwuegbuzie, 2004).
- It allows researchers to adopt a situational approach, namely, to choose appropriate methods for each research question, for optimal solutions; it is argued that mixed methods research present solid evidence for conclusions, by way of providing overlapping multiple evidence (triangulation) confirming the findings (Creswell, 2014; Howe, 2012; Johnson & Onwuegbuzie, 2004).

4.3.1. Introducing data collection phases

As described in Chapter One, the aim of the study is to develop and test a generalisable performance measurement (PM) system for nonprofit organisations (NPOs), using healthcare as the context. Figure 4.1 depicts the strategy (the sequence of fieldwork) implemented by the researcher to answer her research questions.

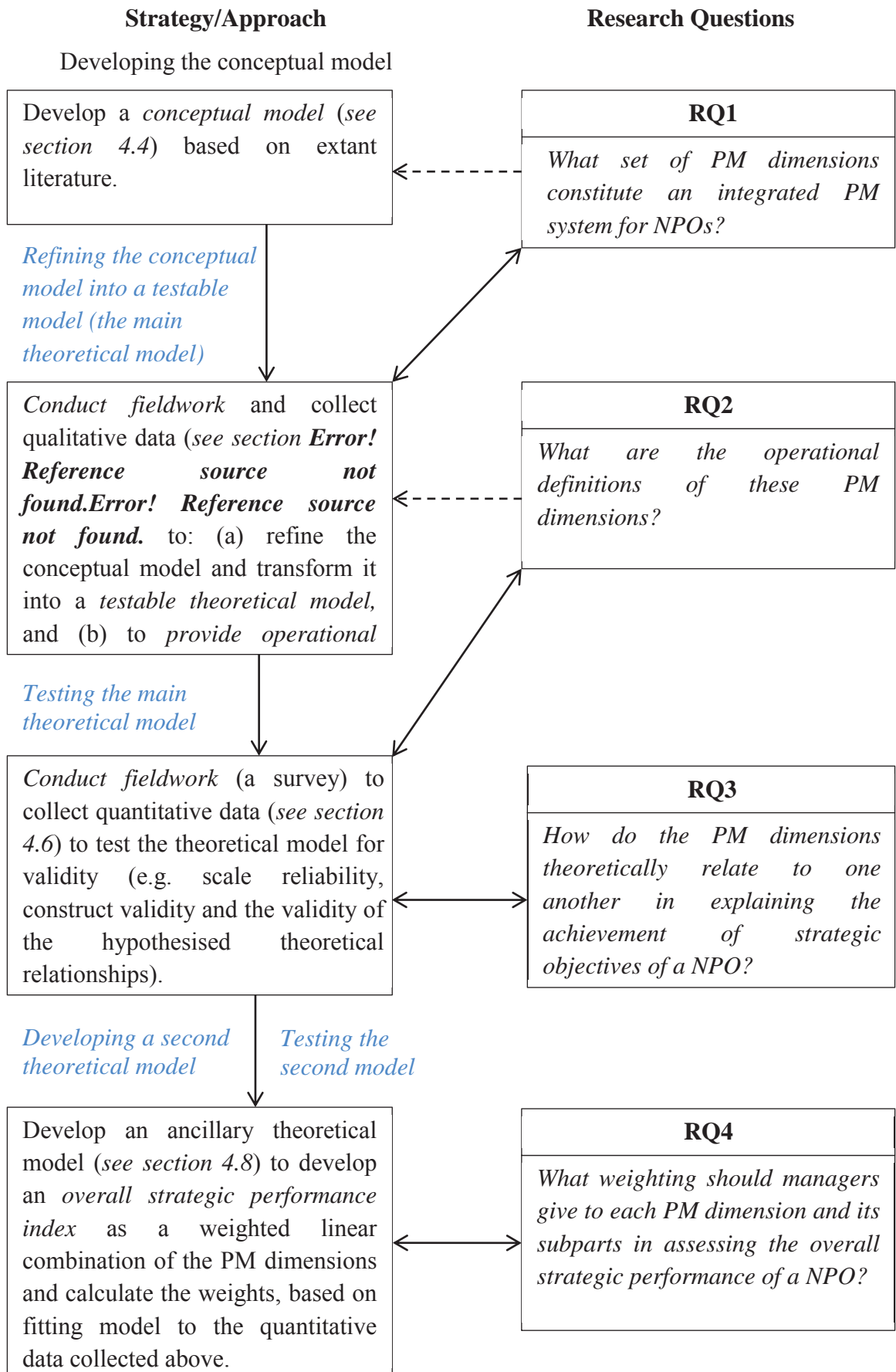


Figure 4.1: The strategy implemented to answer the research questions

As evidenced in Figure 4.1, the researcher builds her models (the main model and the ancillary model) from the literature (the main model was improved later though qualitative data) and tests these using quantitative data. The study, for the most part, involves building hypotheses and testing these for conformity or falsification. Therefore by default, a researcher is tempted to choose a research design that favours hypothetic deductive reasoning espoused in positivism (and post positivism). However, the literature did not provide a definitive set of PM domains (constructs) that underpin the strategic performance of a nonprofit organisation (the literature review led the researcher to posit a conceptual model that is good enough to launch the fieldwork). Equally importantly, there wasn't sufficient *a priori* knowledge (e.g. literature) to operationalise the PM domains (constructs) and to propose testable hypotheses. Therefore refinement of researcher's conceptual model to a testable set of hypotheses—this formed the initial phase of researcher's fieldwork—required interaction with NPOs in a case study environment to understand the phenomena further. This involved the collection of qualitative (non-numerical) data. As evidenced in Figure 4.1, the quantitative data were collected later (using the survey research approach), upon formulation of the hypotheses and operationalisation of the constructs. Thus, the researcher had to adopt a mixed method (in the researcher's case: Qualitative → Quantitative) research design.

Figure 4.2 depicts the flow of the researcher's overall study. A more detailed diagram outlining the researcher's specific mixed method design component is given later (Figure 4.3).

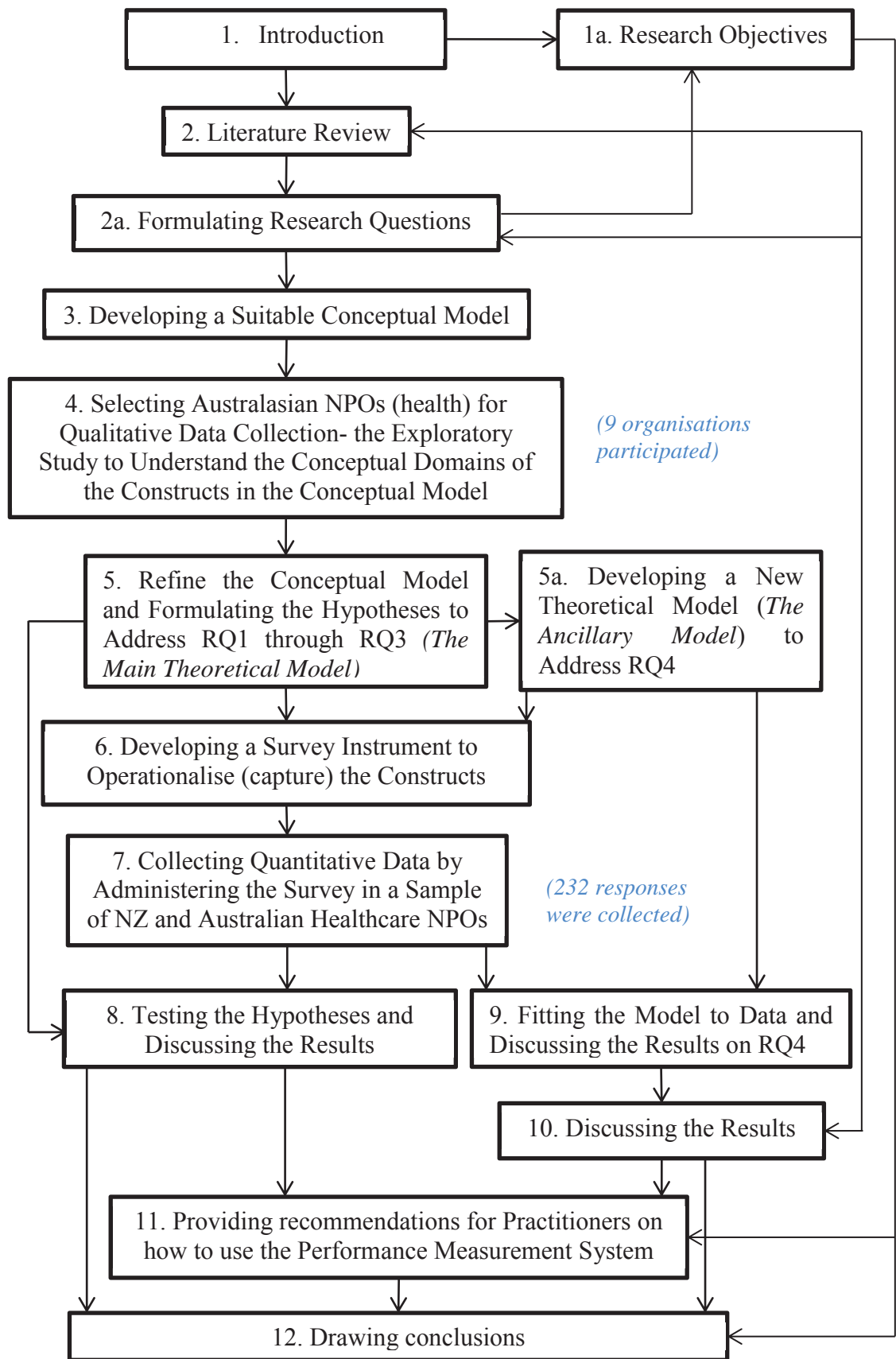


Figure 4.2: Flow of the overall research

According to the literature, there are various design options available to conduct mixed methods research (Tashakkori & Teddlie, 2010). Creswell and Clark (2011) identified six major designs that are commonly used in mixed methods research: *convergent parallel design*, *explanatory sequential design*, *exploratory sequential design*, *embedded design*, *transformative design*, and *multiphase design*.

A convergent parallel design is typically used when a researcher attempts to understand a problem or phenomenon more fully, by collecting and analysing different but convergent (complementary) data concurrently, by merging quantitative and qualitative results (Creswell & Clark, 2011).

An explanatory sequential design is typically used when the purpose of the study is to explain certain quantitative results qualitatively, in which case, quantitative data collection precedes qualitative data collection (Creswell & Clark, 2011).

An exploratory sequential design is typically used when the aim of a study is to identify and measure variables (constructs) that predict or explain a phenomenon and/or to operationalise (e.g. develop a survey instrument) known variables (constructs); in this type of a design, collection of qualitative data precedes quantitative data collection (Creswell & Clark, 2011).

An embedded design (generally regarded as an advanced mixed methods design) is used when the objective of a study is to answer an assortment of research questions that require different types of data (Creswell & Clark, 2011).

A transformative design (an advanced mixed methods design) is a mixed methods design suitable for a study that aims to address issues on social injustice and marginalised groups (Creswell & Clark, 2011).²¹

Finally, a multiphase design (again, an advanced mixed methods design) is a design that achieves an overall programme objective by answering incremental questions (e.g. one

²¹ It seems that a transformative design is tailor made to researchers who subscribe to the participatory/advocacy worldview.

question informing the other) using a combination of sequential and concurrent typologies (Creswell & Clark, 2011).

Out of these six designs briefly described above, the design adopted by the researcher (a variant of the exploratory sequential design), along with the reasons for selection of this design, is described in detail as follows.

4.3.2. Classifying the specific mixed methods design adopted

The instrument development variant of the exploratory sequential design becomes applicable for researcher's study because, to begin with, the qualitative phase provides a supportive role for the quantitative phase; this is the main characteristic of an exploratory sequential design. To elaborate further, the qualitative phase provides the requisite knowledge (based on qualitative data collected from a smaller sample as shown in Figure 4.2) to the researcher on the overall organisational approach towards achieving its Mission (the phenomenon being studied to achieve the research objectives). The qualitative phase was guided by the researcher's conceptual model in framing appropriate questions for her respondents. The qualitative data enabled the researcher to refine her conceptual model to evolve a testable theory, and eventually, build the quantitative tool (a survey instrument) for theory testing, given the sampling frame (Australasian NPOs in healthcare). The reasons for using an exploratory sequential design for the study can be further justified due to the following overlapping reasons:

- The main purpose of the researcher's design is to develop and test a generalisable PM system for Australasian nonprofit organisations (healthcare). The fact of the matter is that there is not adequate prior knowledge out there to accomplish this task. Only a guidance model (a conceptual model) could be developed from prior knowledge (the literature), using the nonprofit BSC as the overarching framework to be augmented. Converting the conceptual model to a testable theoretical model and developing a test instrument (a survey instrument) to measure/capture the constructs of the underlying theory took place only after obtaining qualitative data from a small sample, to understand the constructs in

the first place. This is the hallmark of an exploratory sequential design (e.g. see Morse, 1998).

- Findings of the first phase of researcher's study were used to establish the quantitative method (Greene, Caracelli, & Graham, 1989).
- The researcher's design began qualitatively to facilitate the understanding of otherwise little understood dimensions and constructs of the measurement system that the researcher was aiming to develop (the PM system) (Creswell & Clark, 2011).
- In the researcher's study, exploration of the constructs preceded measurement of the constructs due to inadequate full knowledge on the conceptual domain of the constructs (section 3.4) and the researcher had to discover constructs and develop and test her theoretical model and her test instrument quantitatively. Stated alternatively, model building and instrument development for theory testing remained the end goal of the researcher's approach, which characterises an exploratory sequential design (Creswell & Clark, 2011).
- The researcher's method also allowed her to develop her constructs and the hypotheses progressively to discover theory and the operational definitions of the constructs.
- The method that the researcher used had been adopted previously by other researchers (under the label "exploratory sequential design") to tackle similar problems. For example, Myers and Oetzel (2003) explored the dimensions of organisational assimilation; Mak (2004) examined the constructs of a theory/model on "perceived mattering in young adults' romantic relationships".

There are some methodological similarities between Myers and Oetzel's research (2003) and this researcher's research. Myers and Oetzel used 13 respondents to identify and operationalise—by designing a survey instrument—the underlying concepts (six concepts/constructs emerged) of organisational assimilation (*their first phase of the*

research). This researcher used 9 respondents (from 9 NPOs)—data saturation occurred with 9 NPOs—to refine and operationally define (survey instrument preparation included) the concepts that underpin PM in NPOs (*researcher's first phase of the research*). Myers and Oetzel used 342 responses to validate the constructs (they used confirmatory factor analysis) and to test the relationships between constructs (*their second phase of the research*); this researcher used 232 responses to validate her constructs and to test the relationships between the constructs (*researcher's second phase of the research*).

Both the researcher's approach and the Myers and Oetzels' approach are the same in that while Myers and Oetzel used a confirmatory approach to show the validity of the six constructs that they identified qualitatively (by interviewing 13 respondents), the researcher used a confirmatory approach to show the validity of the nine constructs that she identified qualitatively, through the literature (the conceptual model) and fieldwork (construct refinement and operationalisation). The researcher conducted structural equation modelling (SEM) to validate the constructs and to test the hypotheses on the relationships between the constructs.

The researcher's exploratory research design sequence is illustrated using the template provided by Creswell (2010, p. 58). The reader should note that the researcher finalised two theoretical models: the main theoretical model (Figure 4.6) to explain how the measurement domains (constructs) of the PM system are causally related, and the ancillary model (Figure 4.7) to explain how the overall strategic performance could be derived as a weighted linear combination of the measurement domains (and their subparts) of the PM system.

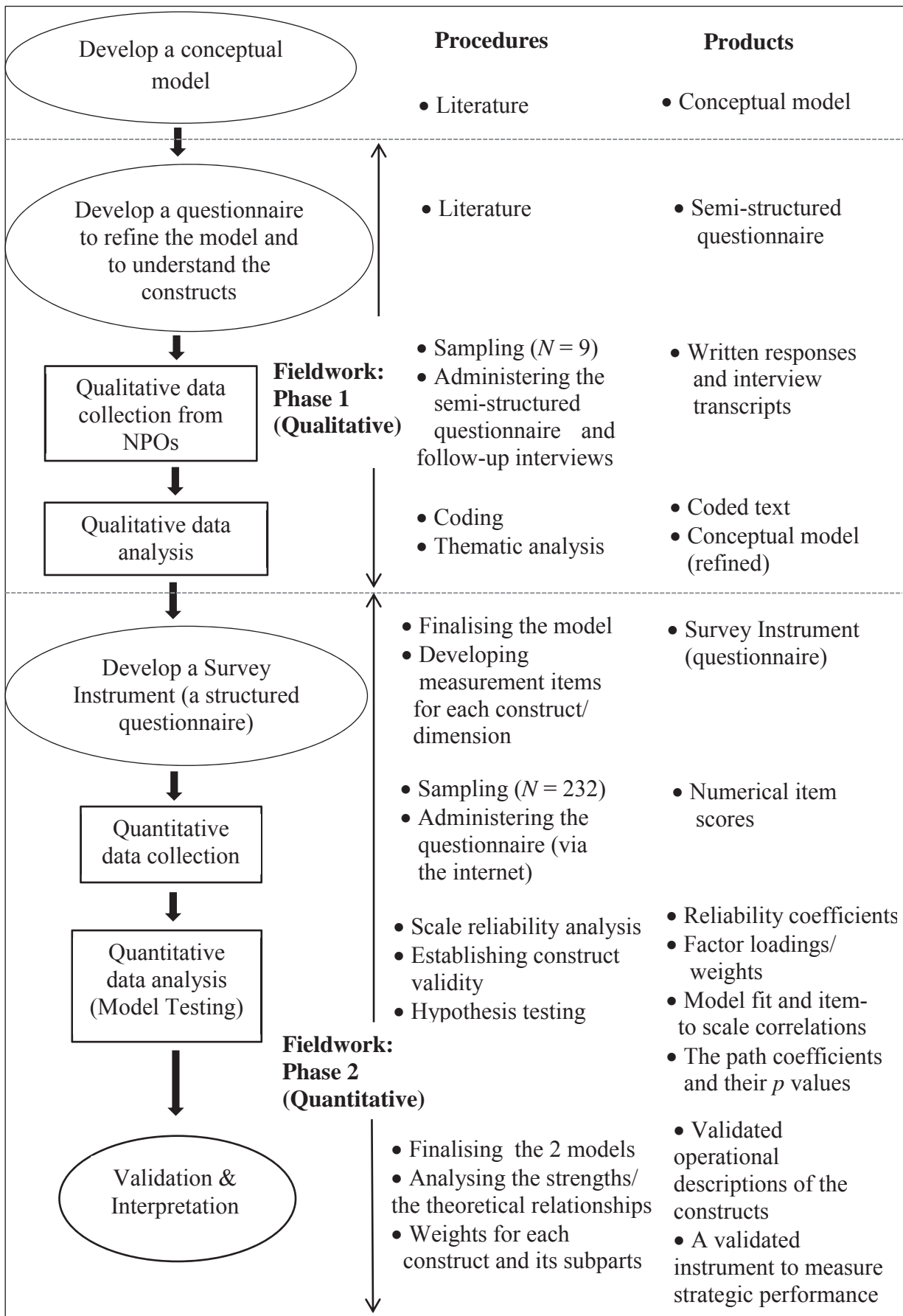


Figure 4.3: The flow of researcher's exploratory sequential design

4.3.3. Justification of the Researcher's Paradigm

The main objectives (the 1st and 2nd research objectives) of this study are to further improve (augment) and operationalise the BSC measurement domains and to develop a scientifically validated PM instrument to test the validity of the measurement domains (constructs) empirically, using the Australasian nonprofit healthcare sector as the sampling frame. The remaining three objectives of the study flow from the first two objectives of the study, mainly through quantitative data analysis.

In order to achieve the key objectives of her study, the researcher has to have a deeper understanding of the concepts that underpin strategic nonprofit PM. However, the literature review did not provide a definitive set of concepts, their operationalisations, and contextual knowledge. Well-defined constructs are vital in developing the hypotheses and measurement models to capture the constructs (these quantitative concepts are explained in the next chapter). This highlighted the importance of understanding what is to be tested and how it should be tested, by interacting with NPOs (qualitative data collection). Eventually, the researcher would adopt hypothetic-deductive reasoning to test her models.

Rather than being locked in a mainstream paradigm (the post-positivist paradigm seems to fit the research objectives because there is nothing that prevents a post-positivist from adopting mixed methods to achieve his/her research objectives), the researcher adopted the pragmatic approach asking herself: *what data collection methods and sequences best help me to achieve my research objectives?* Therefore the researcher's paradigm is 'pragmatism', which focuses on the desired outcomes of the research, not the process.

Prior to undertaking the doctoral study, the researcher worked as a lecturer in engineering mathematics and statistics in a leading university in Sri Lanka. In addition, her prior postgraduate training was on financial mathematics. The researcher realised that this background had a significant impact on shaping the research inquiry.

4.4. DEVELOPING THE CONCEPTUAL MODEL

The conceptual model was developed based on the literature (Chapter Three) on the Balanced Scorecard (BSC) for NPOs (section 3.4.1), the Performance Prism (PP) (section 3.3.3), the Baldrige Excellence Framework (BEF), and the EFQM Excellence Model (EEM) (section 3.3.4). However, the *primary theoretical underpinning* of the conceptual model was derived from the BSC for NPOs.

In Chapter Three, the researcher critically reviewed the literature on key PM frameworks, to examine how they stand in relation to NPOs. The literature review revealed that the nonprofit BSC provides some key elements that are important to management of NPOs (see section 3.4.1 for details). Firstly, it recognises that the Mission of the NPO is what drives the organisation's strategy, and not shareholder objectives such as growth and profitability (Kaplan, 2001, 2008). The BSC also takes a systems perspective in that it recognises that it is the strategic resources and organisational capabilities (Learning and Growth perspective) that enable the Internal Processes to deliver NPO's mission-related outcomes. By induction, it follows that it is the strategy of the NPO that causes the development of its strategic resources and organisational capabilities.

As mentioned elsewhere, Niven (2008) provided more clarity to the nonprofit BSC (please see section 3.4.1 of Chapter Three). In particular, he showed that financial performance (Financial perspective in the BSC) is not a mission-related outcome as envisaged by Kaplan (2001), but a requirement that is needed for developing the intangible resources (Learning and Growth perspective in the BSC). He also argued that in a NPO, financial performance can be improved by taking strategic initiatives on revenue gathering and cost control.

The literature review revealed that lack of full consideration of key stakeholders (e.g. Neely et al., 2003; Yadav & Sagar, 2013), and in particular, overlooking the workforce as a key stakeholder group (hence non-allocation of a dedicated measurement perspective on people-related outcomes) remains one of the major shortcomings of the BSC (e.g. Kong, 2010; Pessanha & Prochnik, 2006). The literature review also revealed that the PP, BEF, and EEM take a more direct stakeholder approach, in that

these models not only take into account the fact that key stakeholder considerations should be taken into account in formulating strategies, but also in assessing process outcomes (details in sections 3.3.3 and 3.3.4). For example, both BEF and EFQM posit that people development is necessary to drive processes (the BSC also does so) to achieve the desired results. Moreover, unlike the BSC, these models also highlight the importance of measuring the actual outcomes on people (intrinsic and extrinsic satisfaction of the workforce) to ensure that the processes do deliver (among other goals) the desired people related outcomes. Managers in a NPO may think that they have taken several initiatives to develop its workforce to drive the processes and satisfy people intrinsically, but it would be hard to conceive that its mission achievement would be complete without its people actually being satisfied. Therefore the researcher argues that process output should cover the satisfaction of three primary stakeholder groups of a NPO: the clients (customers), people (workforce), and the donors.

The above propositions lead to the development of the researcher's conceptual model. This model is depicted in Figure 4.4. The hashed lines imply that being a system, an organisation has (or should have) built-in corrective mechanisms, where preceding inputs receive information from subsequent outputs to iterate towards the desired outcomes. For example, low performance on Strategy coupled with low performance on Mission and Core Values may indicate that the Mission and Core values have not been appropriately transmitted throughout the organisation—thus warranting management action on disseminating Mission and Core Values throughout the organisation.

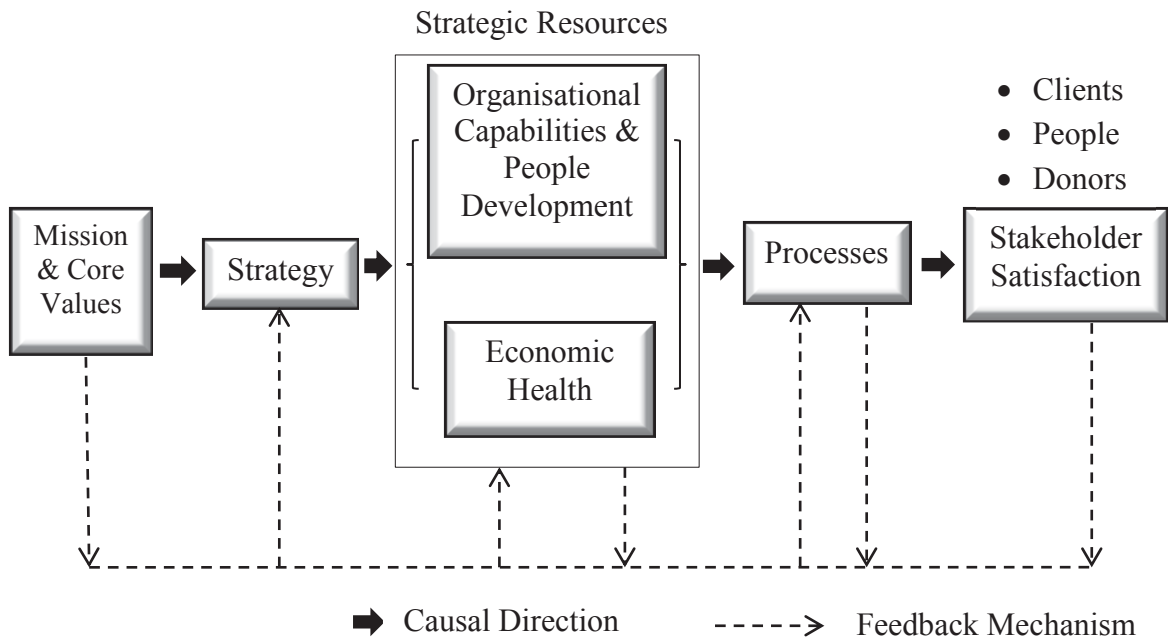


Figure 4.4: The conceptual model developed for Australasian nonprofit healthcare organisations

The core concepts (constructs) of the conceptual model along with operationalisation issues are discussed below. Refinement of the concepts, as mentioned earlier (e.g. see Figure 4.1 through to Figure 4.3), was accomplished through fieldwork (details in section 4.5.1). For the sake of completeness, a brief description of each core concept of the conceptual model is provided below.

- **Mission and Core Values:** The mission of an organisation provides a “sense of purpose” for the organisation to exist (Bryson, 2011, p 102). The mission statement of an organisation is an explicit statement that is meant to motivate/energise its staff in clarifying the question, why the organisation exists (Bryson & Alston, 2004). A closely associated concept with the mission of an organisation is its vision. A vision (statement) clarifies “how the organisation should look like and how it should behave as it fulfils its mission” (Bryson, 2011, p 102). The focus of the researcher, as far as her fieldwork is concerned, is on understanding how the mission of a NPO drives its strategies. This was accomplished by the researcher by tallying the mission statements of NPOs with their corresponding strategic goals and objectives. The researcher has labelled her driver as “Mission and Core Values” to obtain a broader perspective on

Mission. Another important question that needs to be answered through fieldwork is what aspect of mission and core values should be measured and how the mission (and core values) should be operationalised?

- **Strategies:** Strategies (labelled as “Strategy” in the conceptual model) are what an organisation puts in place (e.g. programmes, actions, resource allocations) to realise its mission (Bryson, 2011; Zuckerman, 2012). The conceptual model posits that organisational strategies should be derived according to the mission, vision and core values. The question the senior managers of a NPO should ask themselves is: “what are the strategies that we need to formulate and implement to ensure that the mission is achieved and the stakeholder value is delivered?” One of the two main focuses of the researcher, as far as her fieldwork is concerned, is on understanding how the strategy relates to financial sustainability of a NPO and how a NPO maintains a fit between strategic objectives and its core processes. This is to have a more complete understanding on the causal links between the organisational strategies, financial sustainability (labelled as Economic Health), intangible resources (labelled as Organisational Capabilities and People Development), and core processes (labelled as Processes). The other main focus is in understanding how Strategy could be operationalised.
- **Processes:** In keeping with the systems perspective, the conceptual model views processes as a series of interconnected activities that an organisation (system) undertakes to transform its inputs (financial and intangible resources) into outputs, and thereby, achieve the desired outcomes: Stakeholder Satisfaction (taken as a proxy for *mission achievement*). The question that needs to be answered (in the light of the researcher’s conceptual model), is “what are the processes that have to be put in place in order to allow strategies to be implemented to achieve Stakeholder Satisfaction?” The focus of the researcher is to find answers to this question to operationalise Processes.
- **Organisational Capabilities and People Development:** From a systems perspective, Organisational Capabilities and People Development is what an organisation requires to transform inputs into outputs that result in the desired

outcomes. Organisational capability embraces all the organisational strengths and limitations (employees, technology, skills, knowledge etc.) that enable the organisation to carry out its processes efficiently and effectively, as well as the knowledge the organisation assimilates over the course of its existence. The human resources are the key resource of a NPO; the contribution of people is vital for the sustainability and success of the NPO (it is important to recognise that a sizable portion of the human resource of a NPO may consist of volunteers). Thus the concept Organisational Capabilities and People Development asks the question: “what capabilities should be developed to drive key processes to deliver stakeholder satisfaction?” The focus of the researcher is to find answers to this question to operationalise the concept.

- Economic Sustainability: Even though the financial perspective typically receives low priority in nonprofit performance reporting (because there are no owners or shareholders in a NPO), NPOs still have to have a financial infrastructure to continue serving their customers, develop the organisational infrastructure (e.g. the ICT infrastructure), pay employees (non-volunteers), and meet numerous overheads—in short, maintaining the organisation. One of the two primary focuses of the researcher is to find out how NPOs maintain Economic Sustainability (e.g. to understand the strategies NPOs adopt). The other primary focus is to understand how Economic Sustainability of a NPO should be operationalised.
- Stakeholder Satisfaction: Stakeholder Satisfaction is treated as an abstract concept that is assumed to represent *mission achievement* (the researcher verifies the plausibility of this assumption in her fieldwork), the ultimate goal of a NPO. The primary focus of the researcher, as far as Stakeholder Satisfaction is concerned, is verifying that clients, people and donors constitute the three key stakeholder groups and how these satisfaction domains should be operationalised.

Given that simplicity (without compromising the explanatory power) is a vital characteristic of a good model (Dubin, 1978; Whetten, 1989), consideration was given to reducing the number of constructs, in finalising the conceptual model through

fieldwork. The Mission and Core Values are the starting point in planning strategies (Bryson, 2011). As such, it needs to be ascertained whether the mission and core values are subsumed in strategies. Similarly, it needs to be considered whether Stakeholder Satisfaction (in NPOs) could be equated to achieving the organisational Mission. Conceptually, stakeholder satisfaction in NPOs could be quite different from stakeholder satisfaction in for-profits. In NPOs a large proportion of people may work voluntarily (this is one of the characteristics of a NPO, according to Salamon & Anheier, 1997) and those who donate do not expect a financial return as such. Therefore Stakeholder Satisfaction does not seem to be monetarily driven in NPOs, which challenges the assumption made by Kaplan (2008) in articulating the conceptual foundation of the BSC. Kaplan (2008) argued that Stakeholder Satisfaction is important to a business because without satisfied stakeholders, shareholders cannot achieve optimum long-term returns. The researcher's conceptual model posits that those who provide significant sources of funding (e.g. the government and key donors such as private firms) expect their money will be well utilised by their beneficiaries (the NPOs). Fieldwork corroborated that Donor Satisfaction is a planned strategic outcome of a NPO.

4.5. QUALITATIVE DATA COLLECTION: REFINING THE CONCEPTUAL MODEL

Having developed a conceptual model that underpins PM in Australasian nonprofit healthcare organisations, the next step was to collect data to refine this model. To this end, the researcher adopted a multiple case study approach. The researcher justifies the case study approach because the purpose of refining the conceptual model was to *explore* the rich information pertaining to the underlying constructs of the conceptual model (this enables refining and operationalising the constructs) and refining the explanation of the phenomenon (mission achievement in a NPO) that the model attempts to explain. The constructs of the refined conceptual model can then be treated as dimensions that managers of NPOs need to monitor (and control), to achieve the Mission. Stated alternatively, the constructs of the refined conceptual model are the PM dimensions that constitute an integrated PM system for NPOs. Thus, one purpose of refining the conceptual model was to *answer the first research question (RQ1)*. The second purpose of refining the conceptual model was to gain a deeper understanding

about the PM dimensions to enable their operationalisation, which leads to *answering the second research question (RQ2)*. Qualitative case study data (nonnumeric information collected from NPOs) suits answering these types of research questions (Voss, Tsikriktsis, & Frohlich, 2002; Yin, 2014). The researcher justifies selection of multiple cases because multiple cases enhance external validity and guard against observer bias (Voss et al., 2002; Yin, 2014).

Qualitative data can be collected in many ways (Bryman, 2012). One way is to interview respondents who have the knowledge and experience to provide the data and information the researcher is looking for. Another way to collect data/information is participant observations (ethnography). Interviews are the most common qualitative data collection method (Bryman, 2012; Myers, 2013). In contrast to quantitative data collection via a survey questionnaire, qualitative data collection is aimed at exploring issues deep beneath the surface related to people and organisations (Flick, 2014; Saunders, Lewis, & Thornhill, 2012). Qualitative data for the researcher's study was collected in two steps; first by inviting *written responses* from the respondents to semi-structured questions included in a questionnaire, and then, following up the written responses, with an interview. The data were collected in accordance with Massey University's human ethics guidelines; the entire study, which involves a subsequent quantitative data collection phase, was deemed "low risk" (see Appendix D for the "low risk" ethics approval for the study). The two steps of qualitative data collection are now described.

4.5.1. Sample Selection for the Semi-Structured Questionnaire

Qualitative data were collected from case study organisations (Australasian NPOs in healthcare) using a semi-structured questionnaire containing 22 items (see Appendix A for the complete questionnaire). The questionnaire was answered by a senior manager in each organisation.

In qualitative research, it is important to focus on a particular group (the unit of analysis) in which, all members are similar on such aspects as the occupational level and participant characteristics, as this allows a researcher to explore a particular matter to a greater depth (Saunders et al., 2012). For this reason, a homogeneous purposive

sample consisting of directors, CEOs and seniors executives was selected. The researcher justifies selection of senior managers as the sole respondents because the survey (and the subsequent follow-up interview) relates to strategic issues. These issues are in the domain of a senior manager only (Bryson, 2011; Kaplan & Norton, 2001).

Unlike in probability sampling involving quantitative data, there are no rules of thumb or quantitative procedures (e.g. power analysis) to decide the sample size in case studies involving qualitative data (Myers, 2013; Saunders, 2012). Thus sample sizes was decided based on the ‘information saturation criterion’—that is collecting information until no additional information or new ideas are reached (Saunders, 2012). The sample size determination sequence for qualitative data collection phase is depicted in Figure 4.5.

Although it is desirable to obtain responses from a wider net of respondents of an organisation—top level managers, middle level managers, lower level managers and the frontline workers—awareness (or the lack of it) of strategic issues limited the data collection to managers at the very highest level (this applied to large sample quantitative data collection in the next phase also). The disadvantage of collecting data from top-level managers only is that there is a possibility that the some responses (e.g. on queries related to human resources such as compensation and training; matters related to achievement of social responsibility/client satisfaction) can be biased. Inclusion of multiple cases, interviews, and cross-validation of information through administrative records minimised bias; in the quantitative phase of the study, the bias was tested quantitatively via a method prescribed for determining the absence of bias in quantitative research (section 8.2.2).

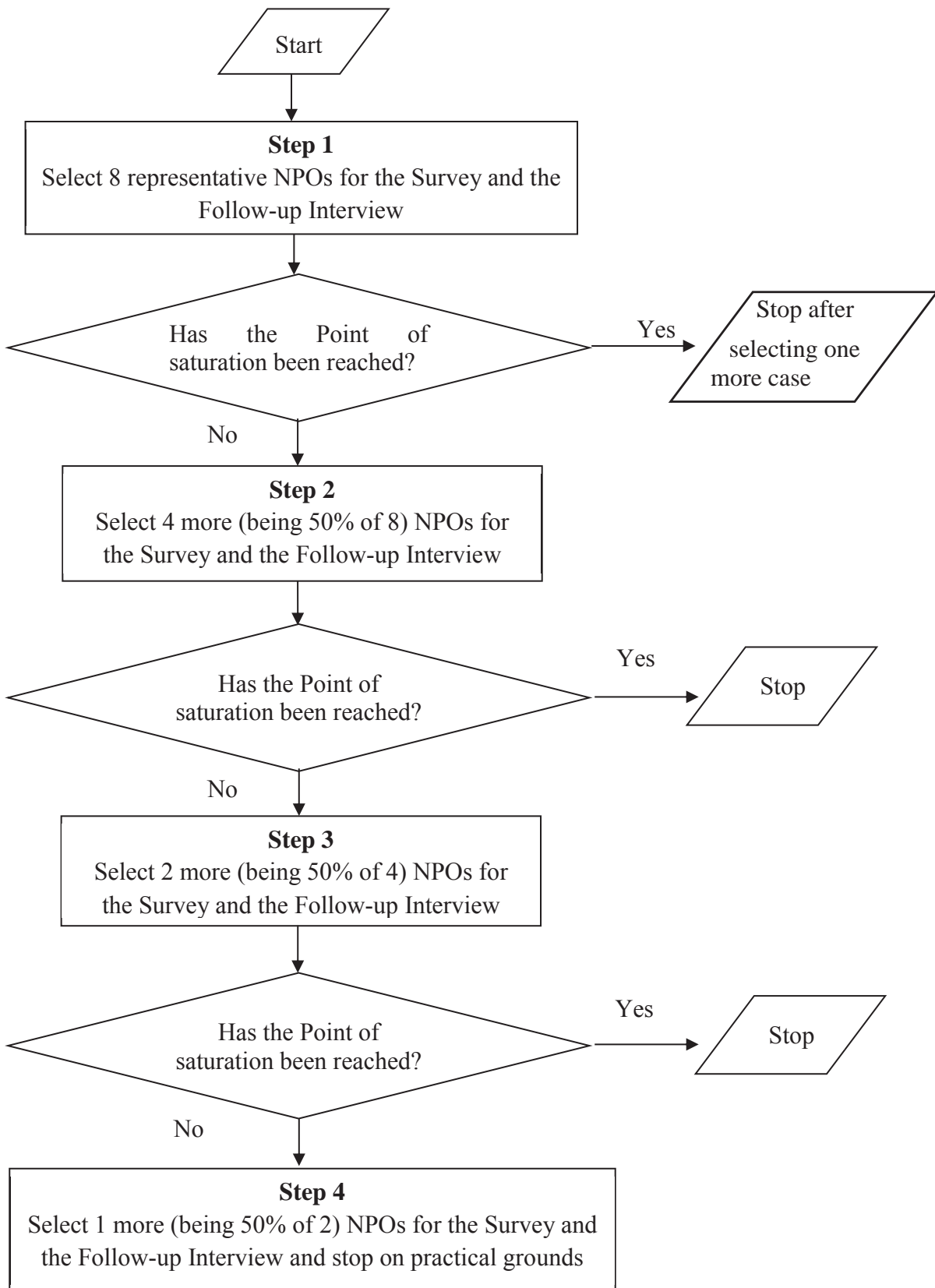


Figure 4.5: The flow chart for sample size determination for qualitative data collection

The outcome of following the sequence depicted in Figure 4.5 was the cessation of collecting new information, after collecting data from nine NPOs. The profile of these NPOs and the designations of the respondents are shown in Table 4.1.

Table 4.1: The Profile of the NPOs and the Respondents

Scope of Operations	Country	No. of Paid Employees	Respondent	Respondent Code
Hospital & Rehabilitation	New Zealand	20	CEO	N1
Nursing Homes	New Zealand	1800	Senior Executive (GM)	N2
Mental Health	New Zealand	450	Senior Executive	N3
Mental Health	New Zealand	42	CEO	N4
Other Health	New Zealand	1500	Senior Executive (GM)	N5
Support Services	New Zealand	1000	Senior Executive (GM)	N6
Injury Care	Australia	127	CEO	A1
Anxiety Prevention	Australia	5	CEO	A2
Women's Health	Australia	1400	Director	A3
Note: A greater representation on New Zealand NPOs is due to logistical reasons (the researcher is based in New Zealand).				

4.5.2. Designing the Semi-Structured Questionnaire

As mentioned earlier, qualitative data collection began by administering a semi-structured questionnaire (Appendix A) on respondents (senior managers as shown in Table 4.1) belonging to a sample consisting of nine Australasian NPOs. Table 4.2 depicts the content of the questionnaire in summary form.

Table 4.2: The Content of the Semi-Structured Questionnaire in Summery Form

Survey Question#	Survey Question
Q1-Q5	Generic questions to justify the selection of the right organisation and the respondent
Q6	Do you have a formal mission statement?; If so, what is it?
Q7	Who are the major stakeholders of your organisation?
Q8	What does each stakeholder expect from your organisation?
Q9	How do you take into account stakeholder expectations in your strategic planning process?

Q10	What progress do you want to see in your organisation in 5+ years' time?
Q11	What high level goals has your organisation established to ensure that it achieves its mission?
Q12	What strategic objectives has your organisation established in order to measure/monitor progress towards its goals?
Q13	What specific strategies do you formulate and implement to ensure that your organisation remains economically sustainable? In particular, how do you ensure that donors/agencies remain interested in providing funding to your organisation?
Q14	List the processes that your organisation has put in place to achieve each of the strategic objectives that you identified in Q12 and Q13 above
Q15	How does your organisation check/ ensure that the strategies that it implements are effective?
Q16	How do you make your paid employees and volunteers aware that their efforts contribute to achieving the strategic objectives of the organisation?
Q17	What are the strengths of your organisation that enables it to achieve its strategic objectives?
Q18	Are you currently using a performance measurement system to measure organisational performance? If so please provide the details
Q19	What are the most important performance measures for your organisation?
Q20	Are you able to identify limitations in some of the performance measures (or the performance measurement system as a whole) that you currently use?; If so what are they?
Q21	How does your organisation fulfil its social responsibility?
Q22	How does your organisation verify/validate that it delivers the expected social impact to the target community?
Note: The generic questions (Q1-Q5) are closed-ended.	

The researcher wishes to mention that the theoretical basis for developing the questionnaire (with the exception of the first 5 questions and question 18) was her conceptual model (Figure 4.4), which, as mentioned earlier, posits that the Mission of a NPO drives its Strategy to enhance Financial Sustainability and Resource Capability to drive the Processes to deliver Stakeholder Satisfaction. Although the theoretical perspective of the researcher has been influenced by the BSC for NPOs, the researcher ensured that she did not directly examine the measurement perspectives of the BSC. This is because this was not one of her research objectives. Rather, as stated earlier, the focus of the researcher was on understanding how a NPO satisfies its stakeholders so that she could formulate a testable theory that explains the phenomenon 'Mission Achievement' (Stakeholder Satisfaction). The constructs of the researcher's theory (model), when fully operationalised, become PM dimensions.

Prior to administering the semi-structured questionnaire, it was pilot-tested²² by requesting a senior executive (the strategic manager) of one of the NPOs that participated in the study, to review the questionnaire for clarity. The said senior executive possessed over 20 years of experience in managing NPOs. As a result of the pilot test, one question in the questionnaire was modified and the final question was removed, as the contents of this question appeared to be repeating a previous question, in a different way. Inasmuch as the researcher would have liked to have used more respondents (NPOs) for pilot testing, limited access to NPOs for qualitative data collection meant that only one NPO was approached for pilot testing the semi-structured questionnaire. After finalising the questionnaire it was dispatched to each respondent via email, in writable 'Adobe Acrobat' format.

As mentioned earlier, the questionnaire-based written response of a respondent was followed up with an interview with the same respondent.²³ The interview was face to face for New Zealand respondents ($N = 6$). For Australian respondents ($N = 3$) it was via the telephone, due to logistical reasons.

4.5.3. The Follow-up Interview

The follow-up interview was necessary for two primary reasons. Firstly, the researcher anticipated that the written responses of some respondents would be too brief (not insightful). Secondly, the researcher knew that she needed to connect the responses with her conceptual model in a meaningful manner, to operationally describe the constructs of her conceptual model, which required filling missing links by asking

²² Prior to data collection, a new questionnaire should be piloted for clarity. More specifically, the researcher should examine each completed questionnaire to make sure that the participants have had no difficulties or the ambiguities of understanding or responding to survey questions (Fink, 2013). This exercise enables in identifying the questions (if any) that require modification/emendation (Fink & Litwin, 2003). Through pilot testing researcher will observe different developments from various fronts or try different approaches on trial basis (Yin, 2014) and identify potential problems before implementing the main study. Furthermore, piloting allows a researcher to make an intuitive judgment on the validity and the likely reliability of the survey questions and the data collection process (Saunders et al., 2012).

²³ It could be argued that the researcher could have approached the respondents and interviewed them straightaway, rather than getting them to answer a questionnaire that contains many unstructured questions, which consumes their time. While this is a reasonable argument, it is important to highlight that the researcher hails from a high power distance culture in the South Asian Region and that English is not her first language. The researcher felt it would be easier for to engage in a conversation with senior-level New Zealand and Australian managers when she has written responses to guide her conversations. This was also an opportunity for to wind her respondents by demonstrating that she is diligent and knowledgeable about what is being questioned.

additional questions related to each questionnaire item, based on what the respondent had provided in writing. The researcher did not employ a separate interview schedule for the follow-up interview. The follow-up interview ranged between 40 minutes to 1 hour, depending on the respondent. The interviews were recorded to facilitate transcribing. In addition to the written responses and interviews, administrative records were also collected, where available and relevant. For example, N6 was a New Zealand Business Excellence Award winner (year: 2013; accolade: gold), and were asked to furnish a copy of their ‘award submission document’. The multiple sources of evidence (written responses, verbal responses, and administrative records) enhanced the quality of the study through comparison techniques and triangulation (Myers, 2013; Voss et al., 2002).

4.5.4. Deriving the Final Model

After conducting the follow-up interviews, these were transcribed into written form for analytical purposes. In doing so, the written response provided by each respondent under each question was typed in (copy & paste) first and the verbal responses transcribed were placed below the written responses initially. Thereafter, the paragraphs (under each question for each respondent) were restructured for coherence to facilitate data analysis. Finally, the information were coded and sorted using ‘Interview Streamliner’ software (Pruijt, 2012). The researcher followed the instructions provided in the coding manual by Saldana (2013) to identify and operationally describe the PM dimensions and sub-dimensions from the textual information that were gathered (details in Chapter Six). As asserted by Saldana (p. 8), coding enables a researcher to “organise and group similarly coded data into categories or families because they share some characteristics—the beginning of a pattern”. This exercise helped the researcher to: (a) operationally describe the constructs, and (b) refine/finalise her conceptual model by modifying the labels of the constructs (where appropriate) and adding new hypothesised relationships between the constructs (details in Chapter Six). The researcher’s final model is shown in Figure 4.6. This model becomes testable when the constructs are operationalised. An important outcome of qualitative data analysis is the separation of the concept Organisational Capabilities and People Development (Figure 4.4) into two concepts/constructs pertaining to strategic resources: People and Organisational Infrastructure (Figure 4.6). It was also evident from respondent information that

Mission and Strategy must continue to be represented as two separate constructs. It was also revealed that mission attractiveness also contributes towards Financial Health (see hypothesis H2 in Figure 4.6). The researcher learnt that donors make donations because they believe in the Mission (donating to a good cause). The details of the findings are given in Chapter Six.

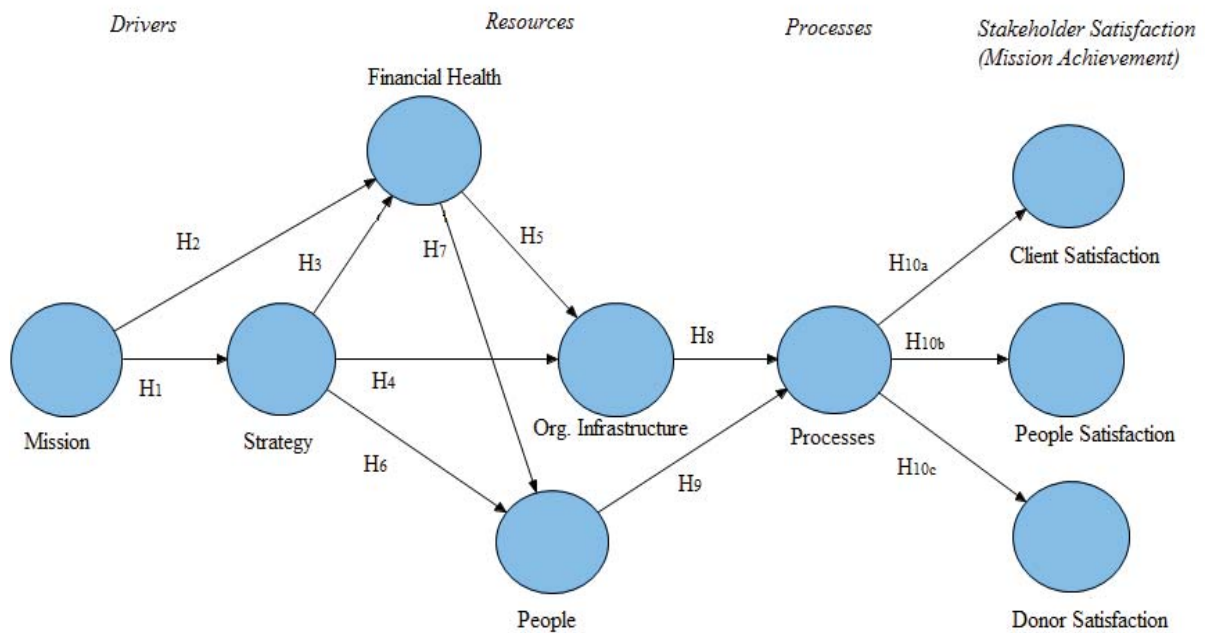


Figure 4.6: The finalised model - the main theoretical model

The following 13 hypotheses were posited for testing

Causal predictors of Strategy

H1: Mission has a positive effect on Strategy.

Causal predictors of the Financial Health

H2: Mission has a positive effect on the Financial Health.

H3: Strategy has a positive effect on the Financial Health.

Causal predictors of the Organisational Infrastructure

H4: Strategy has a positive effect on Organisational Infrastructure.

H5: Financial Health has a positive effect on Organisational Infrastructure.

Causal predictors of the People

H6: Strategy has a positive effect on People.

H7: Financial Health has a positive effect on People.

Causal predictors of the Processes

H8: Organisational Infrastructure has a positive effect on Processes.

H9: People have a positive effect on Processes.

Causal predictors of Stakeholder Satisfaction

H10: Processes have a positive effect on Stakeholder Satisfaction.

The above hypothesis contains three sub-hypotheses:

H10a: Processes have a positive effect on Client Satisfaction.

H10b: Processes have a positive effect on People Satisfaction.

H10c: Processes have a positive effect on Donor Satisfaction.

The above hypotheses provide a tentative explanation of the theoretical relationships between the PM dimensions of the theoretical model (Figure 4.6). Once the hypotheses have been tested, the researcher elaborated on the theoretical relationships (e.g. strong relationships versus weak relationships; new insights the data seem to suggest) to answer her third research question (RQ3).

4.5.5. Operationalising the Constructs of the Final Model

As mentioned earlier, operationalising the constructs of the refined model (the final model) answers the second research question (RQ2). To operationalise the constructs, in the second stage of qualitative data analysis, the coding was arranged at sub-dimension level. The coding process was completed in several steps—data familiarisation, initial coding, theme searching, themes reviewing, themes defining and naming, and reporting—as suggested by Braun and Clarke (2013). Next, similar ideas were grouped together and labelled to reflect the initial coding. With use of the initial codes, the researcher identified the PM dimensions and checked whether these dimensions reflect the frequency and similarity of the responses. This exercise provided sufficient information to operationalise the constructs. In a subsequent chapter (Chapter

Six), the researcher provides the results in the form of operational descriptions of the constructs. The operational descriptions were also transformed into a self-administered survey questionnaire for quantitative data collection (section 4.6).

The number of measures (survey items) found to be necessary to operationalise each construct were as follows:

Mission: 4 measures

Strategy: 5 measures

People: 3 measures

Financial Health: 6 measures²⁴

Organisational Infrastructure: 2 measures

Processes: 8 measures

Client Satisfaction: 4 measures

People Satisfaction: 6 measures

Donor Satisfaction: 3 measures

(Total number of measures = 41; one measure identified under Financial Health was isolated as a global measure for construct validation; details in Chapter Eight)

4.6. QUANTITATIVE DATA COLLECTION

4.6.1. Determining the Sample Size

Apart from selecting the sampling frame, the other most important decision to make prior to quantitative data collection is to know the minimum sample size required to make statistical inferences, without compromising “statistical power” (Cohen, 1992). Statistical power is the probability of rejecting a false null hypothesis (e.g. H_0 : zero effect versus H_A : nonzero effect). Therefore in a regression context, statistical power (which is a function of the sample size, significance level, and the “effect size”) refers to the probability of a regression model parameter found to be significant (based on the p value), if in actual fact (i.e. in the population), it is significant (Cohen, 1992; Green, 1991). In statistical analyses for behavioural sciences, a statistical power of 80% (0.80) is taken as the cut-off value for determining the minimum sample size; a larger sample

²⁴ Unlike the other constructs, this construct was found to be more appropriately modelled as a *formative construct*; details in section 5.2.2.2, in Chapter Five.

size will cause the statistical power to become greater than 80%, if other things remain unchanged (Cohen, 1992; Green, 1991).

The researcher found that the most appropriate statistical method required to test her hypothesised theoretical model (Figure 4.6) is the *partial least squares based structural equation modelling* (PLSBSEM) approach; the justification is given later (in section 4.6.3 and also, in Chapter Five). Chin (1998) asserted that because PLSBSEM is a piece-wise multiple regression method, the minimum sample size required to conduct PLSBSEM should be governed by the multiple regression relationship that contains the highest number of predictors (more technical details are given in section 5.4.2 in Chapter Five). In the researcher's model, the multiple regression relationship that contains the highest number of predictors is the predictor-response relation of the reflective construct Processes. The response Processes is predicted through eight predictors (questionnaire items). Based on the power analysis calculation performed by Cohen (1992, p. 158), the sample size required to attain a statistical power of 80% (at 5% significance level), for a regression model containing 8 predictors, for a medium size relationship²⁵ ("medium effect"), is 107 observations.

4.6.2. Developing the Survey Questionnaire

As mentioned very briefly earlier (section 4.3.1), quantitative data were collected through a self-administered survey questionnaire. This mode of data collection is the most common quantitative data collection method in social research (Bourque, 2003; Bryman, 2012; Cooper & Schindler, 2014; Saunders et al., 2012). The survey questionnaire was administered on representative sets of New Zealand and Australian samples to test the researcher's hypothesised theoretical framework (Figure 4.6); this theoretical framework provides a scientific basis for measuring strategic performance in NPOs. Following are the benefits of adopting a self-administered survey questionnaire:

- It is more efficient, cost effective and an easy way to collect data from a larger sample (Bourque, 2003; Saunders et al., 2012).

²⁵ Cohen (1992, p. 157) uses an arbitrary R^2 value of 13% (more technically, Effect Size (Cohen's f^2) = Explained Variation/Unexplained Variation = $R^2/(1-R^2) = 0.15$) to define a medium size predictive relationship (medium effect), which is actually a weak relationship rather than a medium relationship for most researchers. Stronger the true relationship (effect size), smaller the sample size required to detect a significant model (based on the p value of the F statistic).

- Participants can respond to the questionnaire at their leisure; since top managers are busy most of the time, this type of a questionnaire enables them to assign a suitable time to respond (Bourque, 2003).
- Although self-administered questionnaires are vulnerable to “method bias” (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), there are statistical tests such as *Harman’s Single Factor Test* (Podsakoff & Organ, 1986) to verify that the responses are probably unbiased.²⁶

A *draft questionnaire* containing close-ended response choices was developed to capture each measure belonging to each construct. The questionnaire consisted of two parts. Part one covered the contextual information relevant to the respondent and their organisation while part two covered 41 statements for which agreement was sought on a seven-point Likert type scale; the seven points in the scale corresponded to following agreement levels: strongly disagree (=1), disagree (=2), somewhat disagree (=3), neither disagree nor agree (=4), somewhat agree (=5), agree (= 6), and strongly agree (=7). As mentioned earlier (section 4.5.5), the measures of each construct—in effect the survey instrument—were developed from the fieldwork undertaken during the qualitative (case study) data collection phase of the study. A seven point Likert scale for the questionnaire was used to enhance the reliability of the measurement scales (Nunnally & Bernstein, 1994). Further, the 41 questionnaire items were randomised to minimise method bias (Conway & Lance, 2010; Zikmund, 2013).

The draft questionnaire was prepared using the freeware web interface “Google Forms” (see Google, 2015) for pilot testing the questionnaire as well as for subsequent full-scale data collection. Google Forms is an online questionnaire creation platform that ensures the anonymity of the respondents, which is an important consideration in this study. As mentioned earlier, the survey was conducted in accordance with Massey University’s human ethics guidelines (see Appendix D for the “low risk” ethics approval for the study).

²⁶ The type of bias the researcher refers to here is the “common method bias”. Technically, common method bias can be defined as a systematic measurement error that results in variance in data “that is attributable to the measurement method rather than to the construct of interest” (Podsakoff et al., 2003, p.879).

4.6.3. Selecting the Sampling Frame and Full Scale Data Collection

Positivist research (hypothetic-deductive reasoning) aims to make scientific generalisations using statistical inference techniques (Creswell, 2014). Even though the researcher adopts a pragmatic paradigm, her final goals are hypothetic-deductive in nature (the 13 hypotheses associated with her model were specified earlier). Therefore, it is important to draw the sampling frames (units of analysis) and the respondents (units of measurement) appropriately. In this study, for the purpose of administering the survey, the researcher assumed that it is best if she could focus on a particular group of respondents in the organisation in which all members are similar on such aspects as the occupational level and participant characteristics. For this reason, a homogeneous purposive sample (see Saunders, 2012; Saunders et al., 2012) consisting of directors, CEOs, and seniors executives, was selected as respondents to whom invitations were sent out to participate in the survey. The reason for selecting senior level managers only was because (much like the case of qualitative data collection that preceded) the scope covered in the survey questionnaire relates to strategic management, which is a senior-level management activity.

The researcher invited 775 Australasian healthcare NPOs (the sampling frame) to respond to the questionnaire. In this regard, 465 Australian organisations (healthcare only) were randomly selected²⁷ from the “Charity Registry” maintained by the Australian Charities and Not-for-profits Commission (ACNC), Melbourne, Victoria, Australia (the ACNC registers all Australian NPOs). Similarly, 310 New Zealand organisations (healthcare only) were randomly selected from the “Charities Register” maintained by the department of internal affairs, Wellington, New Zealand. Having selected the 775 organisations, the web sites of all of these organisations were searched to collect the relevant contact information (the name, designation and the email) of senior-level managers (no more than 2 from each organisation) to prepare the list of potential participants in the survey. Thereafter, self-addressed emails were sent through to 1550 (930 to Australia and 620 to NZ) individuals who belong to the 775 organisations (i.e. two senior managers per NPO). The researcher anticipated that there would be at least a 10% response rate, which meant that the anticipated number of responses (≥ 155 responses) would be far greater than the minimum sample size

²⁷ In the random selection process, if a particular organisation was found to have no website, that organisation was excluded.

required for the statistical analysis (see section 4.6.1). The email sent through to the respondents consisted of a brief description about the research; the researcher's background, the human ethics approval, and the link to the web page containing the online questionnaire in "Google Forms" format (see <https://docs.google.com/forms/d/1iCod4I1MLcnmKBCGfbq-V3ZVEhYRkUKmmGqE16BZ8YA/viewform>).

Statistical software packages IBM SPSS 22 and SmartPLS 2 were used to analyse the data. The IBM SPSS 22 software package was used to conduct principal components analysis (PCA) on the 41 survey items of the questionnaire as well as to assess scale reliability (Cronbach's alpha) of the constructs, while the second software package SmartPLS 2 was used to test the hypothesised theoretical model (Figure 4.6) using the PLSBSEM approach (Hair, Hult, Ringle, & Sarstedt, 2014; Jayamaha & Grigg, 2014). The PLSBSEM method (see Chapter Five for details) was used instead of the more widely known covariance-based structural equation modelling (CBSEM) method (incorporated in software packages such as LISREL, AMOS, EQS) due to the following reasons:

- (i) The study involves measurement of new constructs (i.e. constructs that have hitherto not been understood well enough) for which achieving a "covariance fit" (the global goodness of fit pursued in CBSEM) becomes difficult (Chin, 1998; Hair et al., 2014).
- (ii) One construct of the model (Financial Health) contains a formative construct. Since the variability of the measures of a formative construct is not explained by the construct of the theoretical model, the analysis of the covariance structure of the measures in the measurement model becomes problematic, thus making the use of the CBSEM method problematic (Diamantopoulos, Riefler, & Roth, 2008; Roberts & Thatcher, 2009).
- (iii) Some constructs of the model (Figure 4.6) contain a prohibitively large number of measures to use CBSEM. Accommodating a large number of measures in a construct leads to covariance fit issues in CBSEM (in this method, it is

recommended that the number of measures be limited to 5 measures or less, per construct) (Bollen, 1989; Byrne, 2010; Kline, 2011).

PLSBSEM (details in the next chapter) is a nonparametric approach (Chin, 1998). Hence the statistical significance of all model parameters was examined using a resampling technique; SmartPLS 2 uses the resampling technique “bootstrapping” (Efron, 1979) to generate the t values of the parameters, to determine their statistical significance. It is important to note that PLSBSEM assumes that the data (in the researcher’s case survey data coded by the researcher) used in the analysis is interval or ratio type (Hair et al., 2014). However, the Likert scale used in the questionnaire was ordinal (a Likert scale is always ordinal) but the researcher treated that scale as ‘interval’ as other researchers do, in order to conduct PLSBSEM; since there are no known PLSBSEM software packages that specifically handle ordinal data (Hair et al., 2014), a justifiable interval approximation was needed. Under these circumstances, as prescribed in the literature (e.g. Chin, 1998; Hair et al., 2014), ordering of response choices was done to conform to the main requirement of selecting a meaningful midpoint (“neither agree nor disagree”) to make the interval approximation tenable. In addition, at least in theory, opting for a seven point Likert scale in favour of a five-point Likert scale also helped because this made the gaps between response choices narrower (hence approximately similar intervals between response choices).

4.6.4. Tests Conducted Prior to Full Scale Data Collection

Prior to administering the survey full-scale, the following procedures were adopted.

4.6.4.1. Content validity and Face Validity Assessment

The term *content validity* refers to extent to which the measures that have been selected to represent a construct (in essence the content), represent (in a sampling sense) the universe represented by the construct (Nunnally & Bernstein, 1994). *Face validity* on the other hand is a judgement on the operationalisations of the constructs as they look (i.e. on its face) (Nevo, 1985). While both of these seemingly similar forms of validity require judgement, establishing content validity requires more expert judgement than establishing face validity²⁸ (Kline, 2011; Nunnally & Bernstein, 1994). For the purpose

²⁸ This is not to say that the face validity should be judged by a lay person!

of content validity and face validity, the draft survey instrument was provided to four independent academics knowledgeable in PM. The views of four academic experts were solicited to make comments on the questionnaire, in terms of clarity and content. This resulted in minor changes to the wording of two survey items in part two of the questionnaire.

4.6.4.2. Pre-testing and Pilot Testing

Pre-testing basically achieves the same goal that pilot testing achieves. However, pre-testing is a cognitive approach (behaviour observation in a laboratory-like setting) that sometimes offer superior benefits over pilot testing. This is because pre-testing typically creates an environment for the respondents to open-up (Collins, 2003; Groves et al., 2011; Yin, 2014). Inasmuch as the researcher would have liked to have called a few respondents—the CEOs, directors and senior executives—to her research facility to conduct a pre-test, this was deemed not practicable due to the nature of the respondents; typically, pre-testing requires procedures such as requesting each respondent to readout each question loud as they answer it (Collins, 2003). Making such requests to Australasian senior executives is not practicable. Therefore, only pilot testing was conducted as a precursor to the full-scale survey.

For pilot testing, the 9 organisations that participated in the qualitative data collection phase (the case study) of the project were solicited to respond to the survey questionnaire (they were directed to the URL containing the online survey instrument). Each organisation was requested to nominate four respondents to respond to the pilot survey (the idea was to receive about 15-20 responses). Specifically, the respondents were asked to comment on any particular survey item that they found ambiguous. The researcher received 13 responses from 9 organisations. Curiously, none of the respondents suggested changes to any of the survey items (the comments that the researchers received were complimentary). The researcher assumed that by confining the pilot test to the very organisations that provided information (qualitative data) to operationalise the constructs, she could get more critical feedback on her survey instrument.²⁹

²⁹ A complimentary response such as “thank you Ishani the questionnaire looks good to me” is probably not critical feedback!

4.7. QUANTITATIVE DATA ANALYSIS

The researcher received 232 responses (121 from Australia and 111 from New Zealand). Given that the minimum size requirement is 107 observations (as shown in section 4.6.1, a sample size of 107 will attain an 80% statistical power for a medium effect size), a sample size of 232 means a considerably higher chance of detecting true significant effects; statistical power was found to be 98.6%.³⁰ The following analyses were conducted on the survey data.

4.7.1. Data Screening

As the initial step of data analysis, the data were screened to check whether there have been any data entry errors: unusual characters being entered and/or out of range values been entered. Only characters that were acceptable were integers from 1 to 7 for data cells containing responses for survey questions relevant to the 7-point Likert scale. Next, the cells were examined to ascertain whether the missing values have been appropriately represented (this facilitates identification of missing data cells and imputation of data for missing data cells), and how many cases (data records/rows) need to be discarded due to excessive levels of missing values.

Missing data is typical in survey research on social and behavioural sciences (Hair et al., 2014). Usually, a case (a data record) is deleted from the data file if the number of missing values corresponding to the case (data record/row) exceeds 15%. Given that the quantitative survey questionnaire contained 41 survey items in the Likert scale, 15% missing values correspond to six unanswered questions (survey items) by a particular respondent (case). Thus if any case (data record/row) contained seven or more unanswered questions, these were deleted. In addition, if there were to be responses from a respondent (case) who has provided the exact same level of agreement to all survey items (e.g. a 6, meaning “agree” in the Likert scale) such cases were to be treated as “suspect”, warranting “case-wise data elimination”. Perhaps unsurprisingly there was not a single suspect case (a flat score or something close to being flat) warranting deletion based on this criterion. This resulted in the deletion of 9 cases, resulting in a usable sample size of 223. The data cells that contained missing values were filled with

³⁰ The following online power analysis calculator was used: www.danielsooper.com/statcalc3/calc.aspx?id=1

the data imputation option “median of nearby points” in SPSS (Gerber & Finn, 2013, pp 23-24).

After the screening, data were further examined using IBM SPSS 22 software for univariate outliers (any z value > 3.0) and univariate normality (frequency distribution histograms and normal probability plots in the form of Q-Q plots). There weren't any univariate outliers and the data were found to be approximately univariate normal based on the frequency distribution histograms and the Q-Q plots.

4.7.2. Verifying the Absence of Common Method Bias

Podsakoff *et al.* (2003, p.879) define common method bias as a systematic error that results in variance in data “that is attributable to the measurement method rather than to the constructs the measures represent”. The researcher conducted “Harman’s single factor test” (Conway & Lance, 2010; Podsakoff, MacKenzie, & Podsakoff, 2012; Podsakoff *et al.*, 2003) to ensure that the responses probably do not suffer from common method bias. If the responses to a survey suffer from common method bias, according to the Harman’s single factor test, the principal components analysis (PCA) of the survey items would extract only a single factor (component). IBM SPSS 22 software was used to perform the PCA required for Harman’s single factor test. The test results suggested that the responses do not suffer from common method bias. The details are discussed in section 8.2.2.

4.7.3. Establishing Scale Reliability

Scale reliability is a necessary condition for construct validity of a measurement system (measurement scale) (Nunnally & Bernstein, 1994). There are two indicators that are commonly used to test the reliability: Cronbach’s coefficient alpha (Cronbach, 1951) and the composite reliability (CR) (Werts, Linn, & Jöreskog, 1974). Scale reliability measures the degree of relatedness of measures that belong to a construct (Nunnally & Bernstein, 1994); scale reliability is a concept that is only applicable to a traditional (reflective) construct (see Chapter Five); this is because it is only in traditional constructs that the measures are required to co-vary (Diamantopoulos *et al.*, 2008; Jarvis *et al.*, 2003).

Due to the above logic, which is common to all traditional (reflective) constructs, the measure “average variance extracted” (AVE) (Fornell & Larcker, 1981) is also used as an additional measure of scale reliability; the researcher wishes to reiterate that scale reliability is a concept that is applicable to traditional constructs only. The coefficient AVE shows the proportion of variance of the construct that is being shared by its measures, on average (Chin, 1998). As mentioned earlier, the researcher used IBM SPSS 22 software package to test the reliability of the constructs. The CR coefficient and the AVE were obtained after conducting PLSBSEM via SmartPLS 2. The rule of thumb cut-off (minimum) values for Cronbach’s coefficient alpha and CR coefficient (ρ) are 0.70 for a mature or established (mature) construct and 0.60 for an emerging or new construct (Nunnally & Bernstein, 1994; Werts et al., 1974), while the rule of thumb cut-off (minimum) value for AVE is 0.50 (Chin, 1998).

4.7.4. Establishing Construct Validity

In PLSBSEM, construct validity for a *conventional construct* can be determined through convergent validity and discriminant validity (Chin, 1998; Fornell & Larcker, 1981). One way to examine these forms of validities is to observe the loading cross-loading matrix of the measures. The term *loading* refers to the correlation between a measure and its assigned construct; the term *cross-loading* refers to the correlation between a measure and a construct to which the measure is not assigned (Gefen & Straub, 2005). The convergent validity is shown if strong loadings (correlations > 0.70 desired) are observed while discriminant validity is shown when the cross-loadings of a measure are less than its loading (Gefen & Straub, 2005). In PLSBSEM, item loadings and cross-loadings can be easily computed because the PLSBSEM algorithm calculates the factor scores of the constructs based on the specified model.

The literature on PLSBSEM prescribes an alternative way to test the discriminant validity using the “Fornell-Larcker criterion” (Fornell & Larcker, 1981). The discriminant validity is shown when the square root of the AVE of a construct is greater than its correlation with any other construct (Fornell & Larcker, 1981).

4.7.5. Model Testing

In structural equation modelling (SEM), irrespective of the method being used (i.e. PLSBSEM versus CBSEM), it is necessary to make a judgement on the overall model, in terms of its goodness-of-fit to the overall model to the data, before making judgements on the statistical and practical significance of the hypothesised relationships (Figure 4.6) (Chin, 1998; Hair et al., 2014). The reader is reminded that the statistical validity of the hypothesised relationships is required to answer the third research question (RQ3). Unlike in CBSEM, there is no well-established global (overall) goodness of fit measure in PLSBSEM to establish whether or not the model as a whole (the measurement and structural models combined) is a good fit to the data (technical details are in the next chapter).³¹ Thus in PLSBSEM the quality of the overall model is assessed through the quality of the measurement model and the structural model (Chin, 1998; Hair et al., 2014). In PLSBSEM, communalities (or loadings), composite reliability and AVE are used to test the measurement model while the coefficient of determination (R^2), Predictive relevance (Q^2), size and significance of path coefficient, Cohen's f^2 effect sizes, and the q^2 effect sizes are used to test the structural model (Chin, 1998; Hair et al., 2014). Therefore these parameters were assessed (more technical details are found in Chapter Five) to examine the goodness-of-fit of the overall model (results in Chapter Eight).

4.8. DEVELOPING AND TESTING THE ANCILLARY MODEL

The model referred to earlier (Figure 4.6), which the researcher referred to as her *main theoretical model*, pertains to the first three research questions only (see Figure 4.1). The fourth research question (RQ4) asks: “what weighting should managers give to each performance dimension and its subparts in assessing the overall strategic performance of a NPO?”

The nine constructs in the main theoretical model (Figure 4.6) are the PM dimensions that are required to predict and explain Stakeholder Satisfaction (mission achievement). The question then is how could the researcher model a valid Overall Performance Index

³¹ Although modern PLSBSEM software packages report an overall goodness of fit (to be interpreted the same way an R^2 value is interpreted in regression analysis) based on the method prescribed by Tenenhaus, Amato, and Esposito Vinzi (2004), its relevance and the theoretical basis has been questioned by scholars (e.g. see Hair et al., 2014, p.78).

as a weighted linear combination of the constructs of her main theoretical model? Once this is answered, fitting this new model—which, the researcher refers to as her *ancillary model*—to data to estimate the weight of each construct (PM dimension) becomes quite straightforward. This is because the data collected to test the main theoretical model can also be used for estimating the weights.

The hierarchical model that the researcher posited to represent the Overall Performance (as an Index) is shown in Figure 4.7. The purpose of creating this hierarchical model is to fit this model to data to *answer the fourth research question (RQ4)*. The researcher hypothesised that a higher-level abstract concept (Overall Performance) can be formed from lesser abstract concepts (lower-order constructs) that represent different facets of organisational performance (for technical details of hierarchical measurement models, see Chapter Five). The researcher considered two options in deciding on the number of lower order constructs: to use all nine constructs of her main model (option 1) and to reduce the number of constructs through rationalisation (option 2). The researcher opted for option 2, because “parsimony” (simplicity) is a valued property in statistical modelling (Akaike, 1987; Byrne, 2010).³² Accordingly, the three facets of Stakeholder Satisfaction—Client (subdomain 1), People (subdomain 2), and Donor (subdomain 3)—were consolidated as a single concept (Stakeholder Satisfaction). Similarly, the two facets of Capabilities and People Development—Organisational Infrastructure (subdomain 1) and People (subdomain 2)—were consolidated as a single concept (Capabilities and People Development). This resulted in reducing the nine constructs of the main model to six constructs (represented by six middle-tier blue circles in Figure 4.7). The reader will note that the six lower-order constructs almost piggybacks on the six constructs that the researcher utilised in developing her conceptual model (Figure 4.4). The details of parameter estimation (e.g. the λ coefficients representing weights) is given in Chapter Eight (section 8.7).

³² One may ask, why then did the researcher not consider reducing the number of constructs in her main theoretical model (in fact parsimony was considered). In the main theoretical model the researcher is explaining a social phenomenon (how mission of the NPO drives its strategies to achieve stakeholder satisfaction). Therefore the researcher must use the right number of building blocks (constructs) for her explanation (theory). In the ancillary model, the researcher is not interested in explaining a social phenomenon as such. Instead, she is concerned with generating a good representative numerical value for the overall performance. Thus the researcher has the liberty to be more parsimonious.

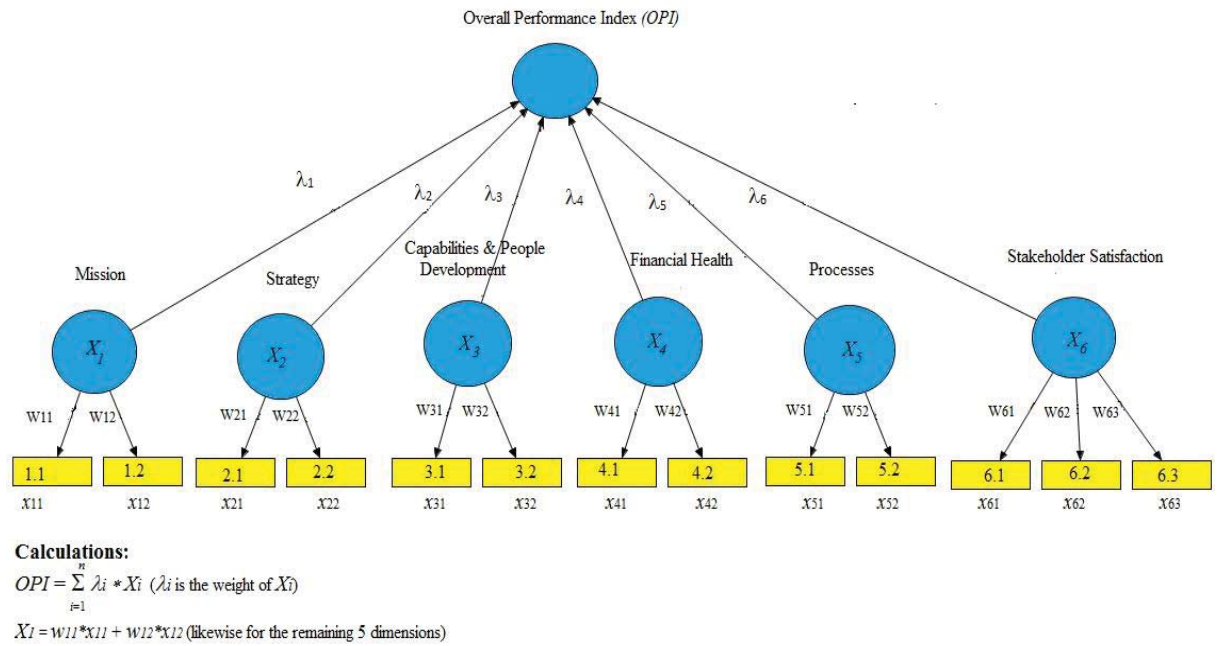


Figure 4.7: The hierarchical model on overall performance

4.9. JUSTIFYING THE QUALITY OF THE RESEARCH

The purpose of this section is to run past specific actions that the researcher took to ensure that she produces a rigorous piece of academic research. The researcher used the criteria³³ prescribed by O'Cathain (2010) to assess the quality of her mixed method research design. O'Cathain prescribed eight generic quality domains (criteria) applicable to mixed method research. The researcher reviewed the criteria and found that six of the criteria (quality domains) apply to her study: Planning Quality, Design Quality, Data Quality, Interpretive Rigour, and Inference Transferability. The key elements of each criterion (quality domain) are summarised by the researcher in Figure 4.8. The details of each domain are discussed in the following six subsections.

³³ The criteria were customised to suit the researcher's scope of work.

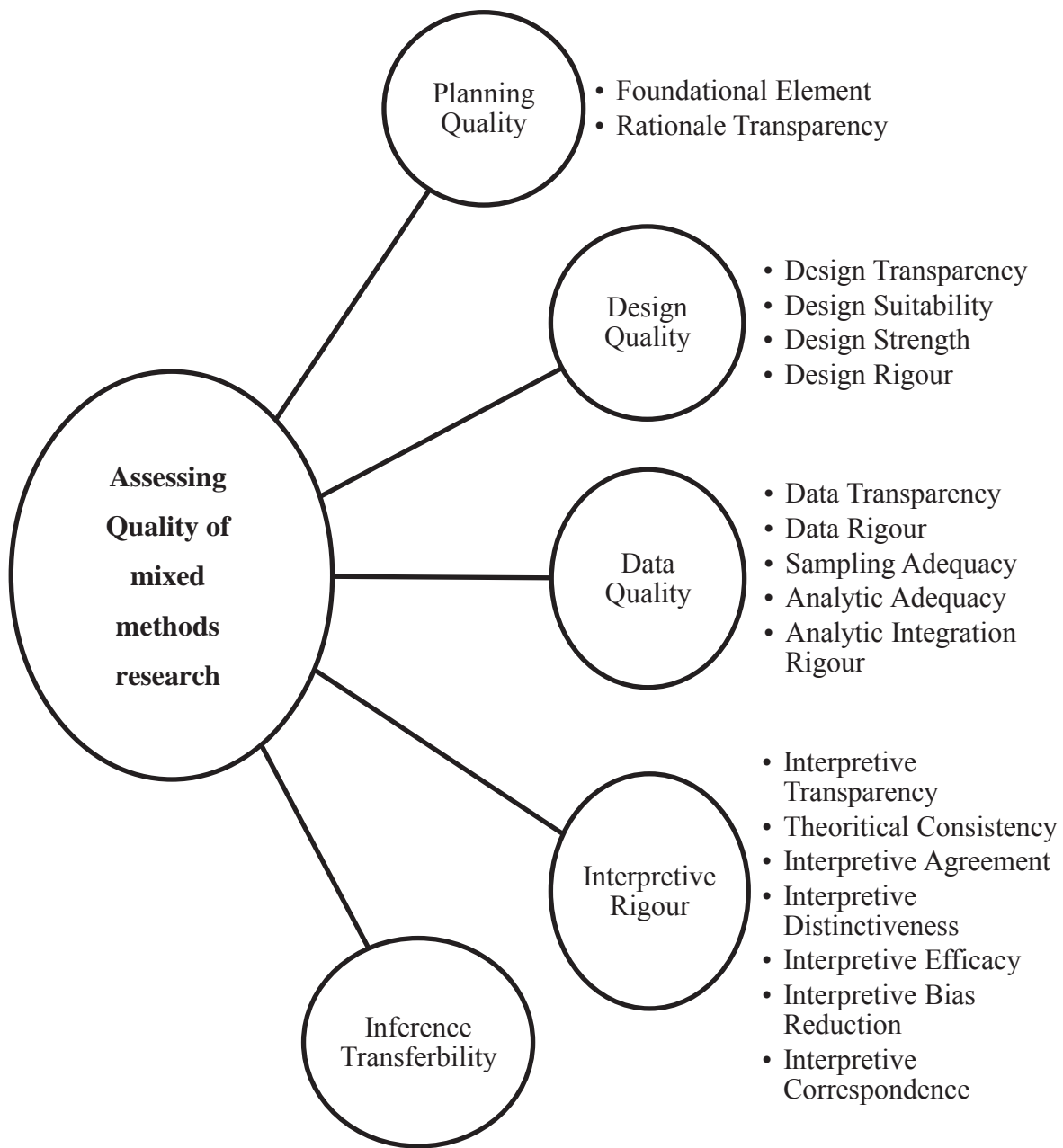


Figure 4.8: The customised quality framework (based on prescriptive criteria given by O'Cathain, 2010) used by the researcher

4.9.1. Planning Quality

According to O’Cathain (2010), two important features have to be addressed to achieve a quality research plan: the ‘Foundational Element’ and ‘Rational Transparency’.

The Foundational Element is a criterion introduced by Dellinger and Leech (2007). This criterion emphasises the importance of understanding and reviewing the existing literature critically, to establish research questions and the study design. Moreover, a conceptual model should be typically developed to guide the design, by way of representing the phenomenon being studied in abstract terms (Caracelli & Riggan, 1994). The researcher followed all these guidelines.

Rational Transparency refers to the justification that is given for adopting mixed methods (Caracelli & Riggan, 1994; Creswell & Clark, 2011); the justification for adopting mixed methods for the researcher’s study was given in detail, in section 4.3.

4.9.2. Design Quality

Design Quality is the criterion that is used to assess the methodological accuracy of a study (Onwuegbuzie & Johnson, 2006). There are four features that are critical in assuring design quality in mixed methods research: ‘Design Transparency’, ‘Design Suitability’, ‘Design Strength’, and ‘Design Rigour’ (Creswell & Clark, 2011; O’Cathain, 2010; Teddlie & Tashakkori, 2009).

Design Transparency is about being transparent in the research design, by relating the design to a known typology or description of important features of the design, if the design does not belong to a known typology (Creswell, 2014; O’Cathain, 2010). The detailed description of the research design selected was provided in section 4.3.2; the specific mixed methods approach (instrument development variant of the exploratory sequential design) used by the researcher was chosen based on the description provided by Creswell and Clark (2011) on mixed methods typologies. For additional transparency, a diagram was provided (Figure 4.3) using the template provided by Creswell and Clark (2011).

Design Suitability refers to the appropriateness of the design to answer the research questions and to provide evidence to justify the stated paradigm and the methods (O’Cathain, 2010). It is almost always the case in published work (especially in high ranked journals) that Design Suitability of the specific mixed methods design is well demonstrated (Creswell & Clark, 2011; Teddlie & Tashakkori, 2009). Onwuegbuzie (2006) used the term “Paradigmatic Mixing Legitimation” to refer to the same requirement. The researcher adopted an Exploratory Sequential Design approach because that approach seemed to be the most suitable mixed methods design type to achieve the aims and objectives of the study (more details were provided in section 4.3.2).

Design Strength provides the evidence of strength of the method (Caracelli & Riggin, 1994; Onwuegbuzie, Bustamante, & Nelson, 2010). In the researcher’s study, the results of the qualitative phase of the study (the case study) were used to develop the PM system (e.g. to refine the conceptual model and operationalise the constructs). These results were then used to develop and strengthen the quantitative phase of the study, which focuses on testing the PM system (see Figure 4.1). Onwuegbuzie (2006) argues that it is possible to optimise a mixed methods design by mixing a weaker approach (weak for the task at hand) with a stronger approach to achieve high quality outcomes, though cross-compensation; Onwuegbuzie termed this approach as “Weakness Minimisation Legitimation”. In the case of the researcher’s study, the qualitative phase of the study (the case study), which led to the finalisation of the main theoretical model, was weak in generalisation (but strong in understanding the meaning of the concepts/constructs of the model); however, once the qualitative phase is launched, it paves way to collect quantitative data to test the model for scientific generalisation (i.e. making inferences about the universe/population to which the study applies). Thus low generalisation of the findings of the qualitative phase (the theoretical model and the operational definitions) is negated by the high generalisation of the findings of the quantitative phase (statistical validity of the model and the operational definitions), presumably resulting in a robust research design. Thus the researcher argues that her design is high in Design Strength.

Design Rigour, the final feature of Design Quality, is the rigour of the methods, given the research design (O’Cathain, 2010). The following actions were taken by the

researcher to assure Design Rigour. These are in line with the recommendations given by Creswell and Clark (Creswell & Clark, 2011) to enhance Design Rigour:

- To identify dimensions of strategic performance (the constructs of the phenomenon being studied) through qualitative coding (phase one) and gain new insights (e.g. new concepts/constructs, relationships, implications) on PM. In phase one, every effort was taken to obtain valid and reliable (trustworthy) qualitative data.
- To develop the survey instrument (quantitative questionnaire) in the second phase to measure otherwise directly unobservable dimensions (i.e. constructs/latent variables) of the conceptual model. Care was taken to develop reliable scales for the constructs (Nunnally & Bernstein, 1994).
- To review the content of the survey instrument by a panel of “content experts” who are knowledgeable about PM in the nonprofit sector; this was to establish face validity and content validity of the survey instrument (content validity is one aspect of construct validity mentioned below).
- To develop an empirically testable theoretical model containing constructs underlying strategic performance and then, demonstrate the validity of these constructs—more technically precisely, construct validity—using established techniques in psychometrics and statistical modelling.
- To examine the similar exploratory sequential design studies to verify that the steps taken by the researcher are constant with peer accepted mixed method designs. As mentioned earlier, the design topologies used by Myers (2003) as well as Mak and Marshall (2004) are very similar to the design topology adopted by the researcher; in addition, studying the methodology used in the peer reviewed work mentioned (albeit their research inquires had very little in common with those of the researcher) also enabled the researcher to become familiar with the specific methods and strategies that are needed to conduct the researcher’s study.

4.9.3. Data Quality

There are several important requirements that have to be addressed by a researcher to demonstrate “Data Quality”.

The first requirement of Data Quality is “Data Transparency”. This is defined as proper description of all the data collection methods, sampling techniques, and data analysis methods (Creswell & Clark, 2011; O’Cathain, 2010). All data collection methods, sampling techniques and data analysis methods that were used by the researcher were clearly described in this chapter.

The second requirement of Data Quality is “Data Rigour”. This is defined as “the extent to which methods are implemented with rigour” (O’Cathain, 2010, p. 546). Onwuegbuzie (2010) used the term “design fidelity” to refer to the same requirement. The action taken by the researcher to ensure Data Rigour, are as follows:

Phase One (qualitative phase/the case study)

- Research records and reports were maintained and updated electronically using Microsoft Word, Microsoft Excel, Adobe Acrobat, Interview Streamliner and Endnote software. In addition, an audio file (where relevant and when possible) of interviews were maintained.
- Action was taken to ensure that the respondents who provide qualitative data are dependable (trustworthy) and that the data are in a format that is suitable for the task at hand.

Phase Two (quantitative phase/theory testing)

- Comprehensive action was taken to develop the survey instrument to test the models (the main theoretical model and the ancillary model) and all methods and techniques that were used to collect, clean (e.g. exclude outliers, impute missing data), and verify data (testing for common method bias) were described in sufficient detail.

- The purpose of quantitative data collection within the overall study was clearly explained.

The third requirement of Data Quality is “Sampling Adequacy”. This refers to the appropriateness of the sample size and techniques (Creswell & Clark, 2011; Onwuegbuzie et al., 2010). Onwuegbuzie (2006) also highlighted the importance of recruiting the right participants in a case study (for qualitative data collection) and combining the two different samples (qualitative and quantitative) in an optimal way to achieve research objectives. This optimal integration of samples is referred to as “Sample Integration Legitimation” (Onwuegbuzie et al., 2010). It is important to note that random selection/assignment of respondents from a universe of respondents is not relevant in a case study, which also means that selection of the cases and the number (sample size) is based on informed judgement, rather than statistical methods (Yin, 2014). This issue was addressed by the researcher in sections 4.3.2 and 4.5.1 (also see Figure 4.5). In the case study, data were collected up until saturation of information was reached (i.e. addition of the last case producing seemingly minimal marginal improvement to the information that was already available to the researcher). The adequacy of the sample size for a quantitative study is typically established by performing a “power analysis” relevant to the statistical analysis to be conducted by an analyst (Cohen, 1992; MacCallum, Browne, & Sugawara, 1996). In the quantitative phase of the study, the researcher determined the minimum sample size required for her study (107 responses as shown in section 4.6.1) based on power analysis and showed that her actual sample size (232 responses) is much greater than the minimum required (see section 4.7)

The fourth requirement of Data Quality is “Analytic Adequacy”. This emphasises the appropriateness and adequacy of the data analysis techniques that were used to answer the research questions (Teddlie & Tashakkori, 2009). The data analysis techniques and the basis of selection of a particular technique (e.g. PLSBSEM) were explained by the researcher comprehensively, particularly in sections 4.5 and 4.7 for qualitative and quantitative data respectively.

The fifth and final requirement of Data Quality is “Analytic Integration Rigour”. This refers to the quality of certain actions that take place during a data analysis process (Onwuegbuzie et al., 2010).

The following were the actions taken to ensure Analytic Integration Rigour:

- Relevant and representative organisations were chosen for data collection in the first phase of the study (qualitative/case study). Only senior executives (e.g. strategic managers) and directors who are actively engaged in performance monitoring were chosen for data collection; this way, the researcher ensured that trustworthy information would be obtained; trustworthiness of informants is critical in qualitative data collection (Creswell & Clark, 2011).
- Relevant and representative organisations were chosen for data collection in the second phase of the study (quantitative/theory testing). Data collection was carried out using an available sampling frame (one for New Zealand and one for Australia). This enabled random selection of Australasian healthcare NPOs to participate in the survey; a random sample makes statistical inferences more meaningful, because statistical inferences are based on the laws of probability (Creswell & Clark, 2011).
- Both qualitative and quantitative strands of data analysis were used to achieve research objectives as it was necessary to do so (Creswell & Clark, 2011).
- The same questions were asked of all the participants in the qualitative phase of the study (the case study) when the researcher invited written responses for her semi-structured questionnaire. Although the researcher did not use a separate interview schedule (not practicable/relevant) in the follow-up interviews, the interviews were restricted (to the extent possible) to seeking clarifications of written responses only. Similarly, a single survey instrument was administered on all potential respondents in the quantitative (theory testing) phase of the study. These steps enhanced reliability and validity of data (Creswell & Clark, 2011).

- The results of the first phase, which resulted in theory refinement and operationalisation of PM constructs, was used to develop the survey instrument (quantitative questionnaire) to test the theory in the second phase, meaning integration of data analytic procedures (Creswell & Clark, 2011).
- Only well-established data analytic techniques were used to analyse the data.

Due to all of the above reasons given in this section, the researcher argues that her study is also high on Data Quality.

4.9.4. Interpretive Rigour

Interpretive Rigour refers to the criteria that are used to draw conclusions from the research findings. There are seven avenues to justify Interpretive Rigour of a study: “Interpretive Transparency”, “Theoretical Consistency”, “Interpretive Agreement”, “Interpretive Distinctiveness”, “Interpretive Efficacy”, “Interpretive Bias Reduction”, and “Interpretive Correspondence” (O’Cathain, 2010). The researcher used all of them!

Interpretive Transparency refers to how the findings emerged from specific data collection and data analytic techniques (O’Cathain, 2010). The researcher has clearly explained how she met this requirement in subsequent chapter sections: section 6.3 (qualitative phase) and section 8.3 (quantitative phase).

The Theoretical Consistency refers to the consistency of the reasoning provided in drawing the conclusions of the study, in the light of existing knowledge or theory (Teddlie & Tashakkori, 2009). Dellinger and Leech (2007) used the term “Inferential Consistency” to establish that the inferences are consistent with existing knowledge. To achieve this objective, the researcher ensured that she discussed the findings in the light of the literature to show how her study contributes to the expansion of the knowledge base on nonprofit PM.

Interpretive Agreement is the consistency of the findings (Teddlie & Tashakkori, 2009). This may be accomplished by meeting the “Inside-Outside Legitimation” requirement proposed by Onwuegbuzie (2010). In Inside-Outside Legitimation, the results would be

reviewed by an outside party and the participants would check the researcher's interpretations of participants' responses. The purpose of this exercise is to eliminate researcher bias. To achieve these objectives, the researcher ensured that the key pieces of her study have been peer reviewed (outside view) through publications; for the inside view, the researcher apprised the participants (in summary form in the quantitative phase) what transpired from quantitative data collection to show each individual that the findings flowed from the responses that they gave.

Interpretive Distinctiveness refers to the degree to which research interpretations are distinct from alternative/rival interpretations and explanations (Teddlie & Tashakkori, 2009). The basic idea in Interpretive Distinctiveness is to provide a strong argument in favour of *internal validity*. The researcher attempted to achieve this objective in two ways. First she posited a theoretical model based on sound available literature and substantive fieldwork (the case study). Second, in the quantitative phase, through testing for common method bias, the researcher ensured that her data have (probably) not been affected by confounding factors.

Interpretive Efficacy, also known as "Political Legitimation" (Onwuegbuzie & Johnson, 2006), is defined as making "meta-inferences from the whole study adequately", incorporating "inferences from the qualitative and quantitative findings" (O'Cathain, 2010, p. 547). The following actions were taken by the researcher to achieve Interpretive Efficacy.

- The conceptual model of the study was developed from the existing literature (e.g. other PM frameworks, systems theory, and multidisciplinary literature). The theoretical basis of the final model and the ancillary model was also explicitly stated. The theoretical basis of the final model is the conceptual model and the case study (qualitative data collected from the case study was used to convert the conceptual model into a testable theoretical model). The theoretical basis of the ancillary model is hypothesising a hierarchical (second-order) model using standard methods used in psychometrics and PLSBSEM.

- Feedback from content experts and pilot tests were used to improve the questionnaires (semi-structured questionnaire and the subsequent online survey questionnaire).
- The results of the quantitative phase of the study (deduction) reinforced the results of the qualitative phase of the study (induction).
- Inferences were made at the end of each phase and after completing both qualitative and quantitative phases, meta-inferences were made and these were peer reviewed (through publications), as advocated by Creswell and Clark (2011).

Interpretive Bias Reduction relates to the reasons given by a researcher for inconsistencies in the findings and inferences (Caracelli & Riggin, 1994; Creswell & Clark, 2011; Teddlie & Tashakkori, 2009). Although the theory testing phase did not produce inconsistencies, the researcher attempted to explain findings from a practical perspective to shed more light on the phenomenon being studied (stakeholder satisfaction/mission achievement in NPOs).

The seventh and final avenue of Interpretive Rigour, Interpretive Correspondence, refers to the level to which the research questions are answered by the researcher (O'Cathain, 2010). Correspondence between research questions and the inferences and meta-inferences drawn from the study have been clearly demonstrated by the researcher, in concluding the research (please see Chapter Nine on conclusions of the study).

For all of the reasons in this section, the researcher argues that her study is also high on Interpretive Rigour.

4.9.5. Inference Transferability

Inference Transferability is the extent to which the study results can be applied to other studies or situations (O'Cathain, 2010). The term “external validity” (generalisability) is used in quantitative research to mean “transferability” — a term used in qualitative research (O'Cathain, 2010; Tashakkori & Teddlie, 2010). One could argue that the

sample selected by the researcher for the qualitative phase of the study leads to results that are not statistically generalisable (case study organisations were not meant to be sampling units). However, this argument can be defended because the quantitative phase of the study did (validation of the measurement model and theory confirmation based on statistical modelling) what the qualitative phase of the study was not meant to do.

To be more specific, in the quantitative phase of the study, the data were collected by administering the survey questionnaire to NPOs drawn from an Australian sampling frame and a New Zealand sampling frame (see section 4.9.6.1) representing healthcare. Having collected quantitative data, the researcher went on to test the validity of the measurement model—that is, the validity of the operationalisations of the constructs (construct validity)—using established techniques in statistical modelling. It is only after establishing the validity of the measurement model did the researcher proceed to test her hypotheses (again using established methods in statistical modelling) for scientific generalisation.

The researcher therefore argues that her study is also high on Inference Transferability.

4.9.6. Other Considerations

4.9.6.1. Samples, Sampling, and the Sample Frames

A group of elements (individuals or organisations) chosen from a larger population for a study is defined as a sample (Bryman, 2012; Cooper & Schindler, 2014). The method or the technique used to choose an appropriate sample from the representative population is defined as sampling (Cooper & Schindler, 2014). A sampling frame on the other hand is a defined collection of sampling units (a list of individuals or organisations that represent the population to which the study attempts to generalise) from which a sample is drawn (Cooper & Schindler, 2014). Two actions are involved in sampling: setting boundaries of the study (based on the research objectives/questions) and creating the sampling frame (Voss et al., 2002).

In the researcher's study the sampling frame is the Australasian NPOs that are classified under the category "health". Both Australia and New Zealand follow the "International

Classification of NPOs” (ICNPO), in which NPOs are classified under 12 main activity groups; “health” is one of them (Salamon & Anheier, 1997).

The New Zealand sample was drawn from a population consisting of NPOs that provide community healthcare from each of the five following subcategories defined by the New Zealand Ministry of Social Development, New Zealand:

1. Hospitals and rehabilitation
2. Nursing homes
3. Mental health and crisis intervention
4. Other health services
5. Support and ancillary services

The above categories fit into NPOs in healthcare, based on the ICNPO.

The Australian sample was drawn from a population consisting of NPOs that provide community healthcare from each of the four following subcategories defined by the Australian Charities and Not-for-profits Commission. These four subcategories are as follows:

1. Hospitals and rehabilitation
2. Nursing homes
3. Mental health and crisis intervention
4. Other health services

The above categories also fit into NPOs in healthcare, based on the ICNPO.

In this study, suitable sampling techniques and adequate sample sizes were selected in each phase of the study. The sampling techniques and strategies that were used in the qualitative phase of the study (the case study) was covered in section 4.5.1 while the sampling techniques and strategies that were used in the quantitative phase of the study (theory testing) was covered in sections 4.6.1 and 4.6.3. Therefore the researcher argues that she completed the sample selection process diligently.

4.9.6.2. Ethical Considerations

The researcher took the following action to ensure that her study conforms to ethical protocols (explicit and implicit):

- As stated earlier (sections 4.5.1 and 4.6.2), Massey University Human Ethics Committee approval was obtained to conduct the study as a low risk study (as mentioned earlier, the research objectives did not warrant a full ethics approval).
- The anonymity of the respondents and their organisations were preserved and confidentiality of strategic information of individual organisations and information of a sensitive nature were not disclosed in any research output, including this thesis (sensitive information is not deliberately sought by the researcher anyway, as such information is not necessary for the researcher to achieve her research objectives).
- The interviews were recorded with participants' permission (as mentioned earlier, this was done to facilitate transcribing).
- The semi-structured questionnaire was dispatched to the participants to invite written responses to each questionnaire item. The respondents were assured that the interviews that followed-up their written responses were meant to seek additional certification of their written responses. Thus in effect, the semi-structured questionnaire served as an interview protocol that defines the parameters within which the interviewer (researcher) operated.

4.10. CHAPTER SUMMARY AND CONCLUSION

This chapter covered the methodology used to answer the research questions, and thereby, achieve the specific research objectives set out at the beginning of the study (see section 1.5). Key social research paradigms were examined (section 4.2) and the appropriate paradigm to answer the research questions was chosen (with justification) as “pragmatism”, which involves mixed methods research approaches (section 4.3).

The first research question (RQ1), which relates to identifying appropriate PM dimensions for monitoring strategic performance, was answered by developing a conceptual model from the literature (section 4.4) and improving it to become a testable theoretical model by using *qualitative data* collected from nine Australasian NPOs in healthcare (details in section 4.5). This fieldwork also enabled the researcher to operationalise the PM dimensions. Subsequent fieldwork of a *quantitative nature* was undertaken to test the operationalisations of the PM dimensions and thereby, fully answer the second research question (RQ2). The quantitative phase of the study (details in sections 4.6 and 4.7) involved data collected from 232 senior-level managers belonging to Australasian healthcare NPOs; data were collected through an online survey. One thousand five hundred and fifty (1550) managers belonging to 775 NPOs (two managers per NPO) were invited to participate in the survey. The quantitative data were also used to test the hypothesised theoretical relationships of the model and discuss their significance, to fully answer the third research question (RQ3). The final research question (RQ4), which relates to determining the weights each PM dimension (and its subparts) should carry to indicate the overall organisational performance, was answered by theorising and testing an appropriate hierarchical factor model (details in section 4.8). Figure 4.1 depicts the strategy implemented by the researcher to answer her research questions, in summary form. Finally, the researcher demonstrated that her mixed methods research design meets the quality criteria prescribed in the literature (section 4.9).

CHAPTER FIVE

REVIEW OF POSITIVISTIC TECHNIQUES USED IN TESTING THEORETICAL MODELS

Do the right thing. It will gratify some people and astonish the rest.

Mark Twain

5.1. INTRODUCTION

As mentioned in the previous chapter, the researcher used quantitative research techniques to test her theoretical models. To enable this, the researcher operationalised the constructs of her models, using case study research. These are positivistic notions of advancing knowledge. Quantitative positivistic approaches in social science research have their own set of quality criteria to assess the rigour of a study (Bryman, 2012; Easterby-Smith et al., 2012). Even though the researcher adopted a pragmatic paradigm to achieve her research objectives, because of the strong quantitative focus in her mixed methods research design, the researcher used certain positivistic techniques, which are reviewed in this chapter. The purpose of this chapter is to review potential quantitative approaches available for testing the validity of theoretical models and to better inform the reader about the quantitative techniques associated with the methodology used in the study, prior to interpretation of the results reported and discussed in subsequent chapters. In a sense, this chapter can be viewed as a chapter on the review of the literature on potential quantitative research techniques available to test the researcher's theoretical models.

5.2. VALIDITY OF A THEORETICAL MODEL FROM A QUANTITATIVE RESEARCHER'S STANDPOINT

5.2.1. The Building Blocks of a Theoretical Model

Constructs, as the name implies, are the building blocks of a theoretical representation. Constructs are *abstract concepts* that a researcher defines for the purpose of theory building (Cooper & Schindler, 2014). A construct is measured by combining more concrete ideas (or simpler concepts) that are directly observable (Cooper & Schindler,

2014; Dubin, 1978). A phrase that goes hand in hand with a construct is the *operational definition of the construct* (Dubin, 1978). An operational definition of a construct is the “set of procedures” (descriptions) that are being used to “measure and manipulate” the abstract concept (i.e. the construct) under observation (Hoyle, Harris, & Judd, 2002, p.11). Because of their abstract nature, constructs are also referred to as *latent variables, factors, or theoretical variables*.

Indicators are the concrete ideas (measures) that a researcher uses to operationalise (capture) the construct. Thus by definition, indicators are directly observable. Indicators are also referred to as *test items, measures, observable variables, or manifest variables*.

The level of measurement refers to the particular way a measurement concerning a variable is treated. The measurement (thus by default an observable variable) can be presented in one of the following four levels or scales: nominal, ordinal, interval, and ratio. A *nominal* scale classifies objects (or test subjects) into mutually exclusive categories (e.g. males versus females, for-profit versus nonprofit). An *ordinal* scale classifies objects/test subjects in a rank-order, but the distance (interval) between the ranks does not convey any numerical meaning (e.g. rating of an outcome as poor, moderate, fair, and excellent); an *interval* scale on the other hand places equal distance between the rank orders, but there is no absolute zero or an origin to the measurement; a ratio scale on the other hand has an absolute zero, which makes it possible to compute ratios between two measurements (e.g. \$ 200 versus \$ 100). The level of measurement concerning a variable is important to a researcher because this determines the appropriate statistical method that should be used by the researcher to analyse (or present) the data (Cooper & Schindler, 2014; Stevens, 1946; Velleman & Wilkinson, 1993).

The *Likert scale* is one of the most widely used attitudinal scales, where the respondents are asked to indicate their level of agreement—ranging from “strongly disagree” to “strongly agree”, scaled typically in four, five or seven points—to a statement given in a questionnaire (Norman, 2010). It is widely accepted in the psychometric literature that attitudinal scales pertaining to a single questionnaire item is ordinal, but when individual questionnaire items are combined to form a single psychometric scale to

measure a latent variable (or construct), attitudinal data can be treated as interval data, to permit the use of parametric statistical methods such as regression analysis and structural equation modelling (SEM) (Boone & Boone, 2012; Norman, 2010; Velleman & Wilkinson, 1993).

A *hypothesis* is a speculative, empirically falsifiable statement that predicts or explains the relationship between two or more variables (Cooper & Schindler, 2014). A term that is closely associated with a hypothesis is a *proposition*. A proposition is also a statement about relationships between abstract concepts, that can be judged to be true or not. A proposition becomes a hypothesis when the proposition becomes testable empirically (Cooper & Schindler, 2014; Whetten, 1989).

A *theory* is a set of interrelated hypotheses proposed to “explain or predict phenomena” of interest (Cooper & Schindler, 2014, p.668). Cooper and Schindler (2014) observe that, in management, many authors use term “model” and theory interchangeably. Cooper and Schindler assert that the two terms are not the same in that the theory provides the “explanation” of the phenomenon whereas the model provides the “representation” of that explanation.

5.2.2. Modelling of Constructs

Constructs need to be represented (modelled) in a particular way to invite mathematical and statistical manipulation, for the purpose of testing hypotheses (Hair et al., 2014; Jarvis et al., 2003; Nunnally & Bernstein, 1994). From the review of the psychometric literature, the researcher identified three ways in which a construct can be modelled, depending on the theorist’s *epistemic* stance on the abstract concept under observation (Chin, 1998; Hair et al., 2014). These three ways are reviewed and discussed in the next subsections.

5.2.2.1. The Standard Form of Modelling a Construct

The *classical test theory* provides the basis of modelling a conventional construct, which is also known as a *reflective construct* (McDonald, 2013). The classical test theory is based on the premise that any measurement of a variable contains its true score, disturbed by an error component that is unique to the variable (Nunnally &

Bernstein, 1994). This assumption is used in developing the measurement scale of a construct. The other assumption on the construct is that a construct embraces a certain surplus meaning to the meanings embodied in the measures (manifest variables) that are being used to develop the measurement scale of the construct (Cohen & Swerdlik, 2002; Jarvis et al., 2003). The implication on this is that any weighted sum of scores of the indicators is, at best, an approximation of the score of their underlying construct (Bollen, 2002; Grice, 2001; Velicer & Jackson, 1990).

A construct, as an abstract concept (a latent variable), is shown either as circle or an ellipse. Figure 5.1 depicts a research hypothesis where the first construct (η_1) explains/predicts the second construct (η_2). The first construct is associated with n number of manifest variables while the second construct is associated with m number of manifest variables. The variable labelled 'R' in Figure 5.1 represents the *residual error*. This variable is also regraded as a latent variable because R corresponds to a difference between the scores of two latent variables, as evidenced in equation 5.3.

The first two equations shown³⁴ (equations 5.1 and 5.2) represent the relationship between the constructs (latent variables) and their associated measures. The parameters λ_{1i} and λ_{2j} represent the association (correlation) between the measures and their corresponding latent variables; these correlations are known as the factor loadings or simply, loadings in the literature (Chin, 1998; Hair et al., 2014). The terms ε_{1i} and ε_{2j} in equations 5.1 and 5.2 the error terms associated with the specified relationships. In statistical modelling involving latent variables, the equations 5.1 and 5.2 are said to represent the *measurement model* (also sometimes known as the inner model) of the overall statistical model (Chin, 1998; Hair et al., 2014). The final equation (equation 5.3) represents the research hypothesis involving the two latent variables. In statistical modelling involving latent variables, the latter equation is said to represent the *structural model* (also sometimes known as the outer model) of the overall statistical model (Chin, 1998; Hair et al., 2014). In a more complex theory, a researcher would specify several structural relationships³⁵ involving constructs, to represent multiple

³⁴ Typically these equations are represented in matrix format but for the ease of understanding the researcher has avoided the use of the matrix notation.

³⁵ Again, it is customary to show all the structural relationships through a single equation, in matrix format.

hypotheses. Moreover, the unknown model parameters (e.g. λ_{1i} , λ_{2j} , and k) are estimated for standardised scores to facilitate the comparison of parameters.

$$y_{1i} = \lambda_{1i} * \eta_1 + \varepsilon_{1i} \quad (i = 1, 2, \dots, n) \quad (5.1)$$

$$y_{2j} = \lambda_{2j} * \eta_2 + \varepsilon_{2j} \quad (j = 1, 2, \dots, m) \quad (5.2)$$

$$\eta_2 = k * \eta_1 + R \quad (5.3)$$

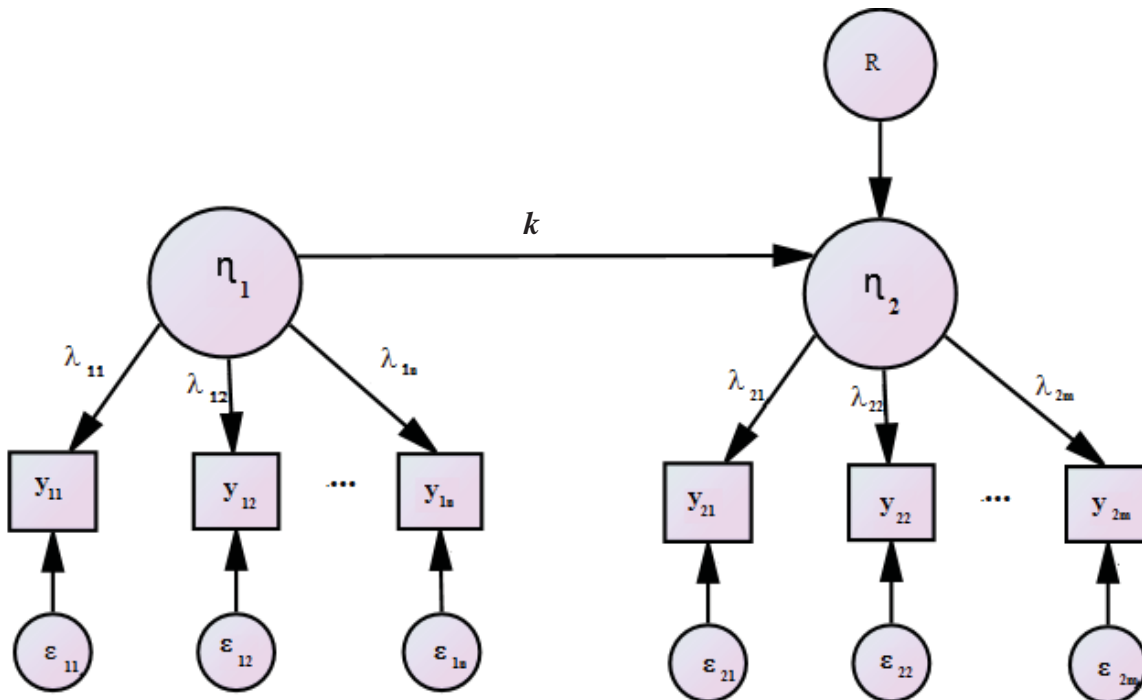


Figure 5.1: Specifying the measurement model and the structural model in graphical format

An important conceptualisation in statistical modelling involving latent variables is that the measures of a construct are *conventionally* being viewed as reflections (manifestations) of the construct, rather than as elements that form the construct (Bollen & Lennox, 1991; Jarvis et al., 2003). Stated in another way, the cause of the measures is their underlying construct (Bollen, 2002; Jarvis et al., 2003). Therefore, when the latent variable varies, all its measures are expected to vary in the same way (Bollen, 2002; Nunnally & Bernstein, 1994).

5.2.2.2. Modelling Formative Constructs

Although a construct is conventionally being viewed as the cause of its indicators, a construct is modelled as the effect of its measures in certain conceptualisations; stated alternatively; a construct can be viewed as being formed by its measures (Fornell & Bookstein, 1982; Gefen, Straub, & Boudreau, 2000). Such constructs are known as formative constructs. Psychometricians assert that the choice between a formative model specification and a reflective model specification should be based on the epistemic considerations involving the relationships between the indicators and the specific construct (Bollen, 1989; Roberts & Thatcher, 2009).

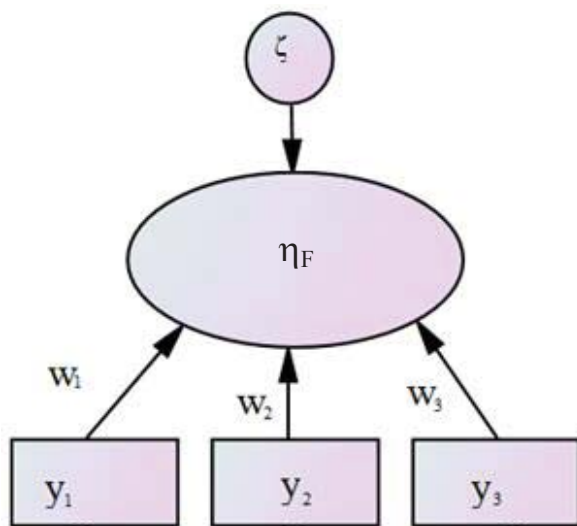
Jarvis et al. (2003) show that the key statistical properties of formative indicators (measures) are different from those of reflective indicators (measures); whereas reflective indicators move in the same direction (because the variability of these indicators is caused by the construct), the formative indicators need not necessarily move in the same direction, because the variability of formative indicators is not caused by the construct (in fact it is exactly the reverse). For this reason, it is known that conventional procedure used to assess reliability and validity of a conventional construct is not suitable to assess reliability and validity of a formative construct (Bollen, 1989). Jarvis et al. (2003) prescribe the following situations that warrant a formative construct:

The indicators are defining features of the construct; changes in the indicators are expected to cause changes in the construct; changes in the construct are not expected to cause changes in the indicators; the indicators do not essentially share a common theme; excluding an indicator may adjust the conceptual domain of the construct; a change in the value of one of the indicators is not essentially expected to be related with a change in all of the other indicators; and the indicators are not expected to have the same causal antecedents and consequences (Jarvis et al., 2003, p.203).

The graphical and mathematical representations of a formative construct (η_F) are provided in Figure 5.2 and Equation 5.4 respectively.

$$\eta_F = w_i * y_i + \zeta_i \quad (i = 1,2,3) \quad (5.4)$$

Where, y_i is the i^{th} indicator of the construct η_F , ζ is the disturbance term associated with the latent variable η_F (it means that the indicators only map partially into the construct), and w_i are the factor weights.



Notes:

1. In this example, the construct (labelled as η_F) is formed by three indicators (y_1 , y_2 , and y_3).
2. The disturbance term ζ is not always included in modelling. When the disturbance term ζ is excluded, that means that the indicators form the construct completely.
3. In some forms of statistical modelling, the researcher has to account for the covariance between the indicators.

Figure 5.2: Illustration of the measurement model of a formative construct

5.2.2.3. Modelling Hierarchical Constructs

In the literature on causal modelling involving latent variables, on occasions, one comes across constructs that represent very highly abstract concepts that require less abstract concepts to operationalise the construct (Chin & Gopal, 1995; Hair et al., 2014). In such situations a hierarchical relationship between constructs is formed, where the less abstract concepts (the first-order constructs) form the basis of the highly abstract concept (the second-order construct), in a reflective or a formative fashion (Chin & Gopal, 1995; Jarvis et al., 2003).

Based on the direction of the relationships between the indicators (manifest variables or measures) and their corresponding first-order constructs and the relationships between the first order constructs and the second-order construct, a second-order construct can be modelled in several different formats (Wetzels, Odekerken-Schröder, & Van Oppen, 2009). Figure 5.3 depicts (in four separate panels) four possible formats of a second-order construct. According to the literature, the approach used in validating a second-

order model is no different to that being followed in validating a first-order model (Chin, 1999; Hair et al., 2014).

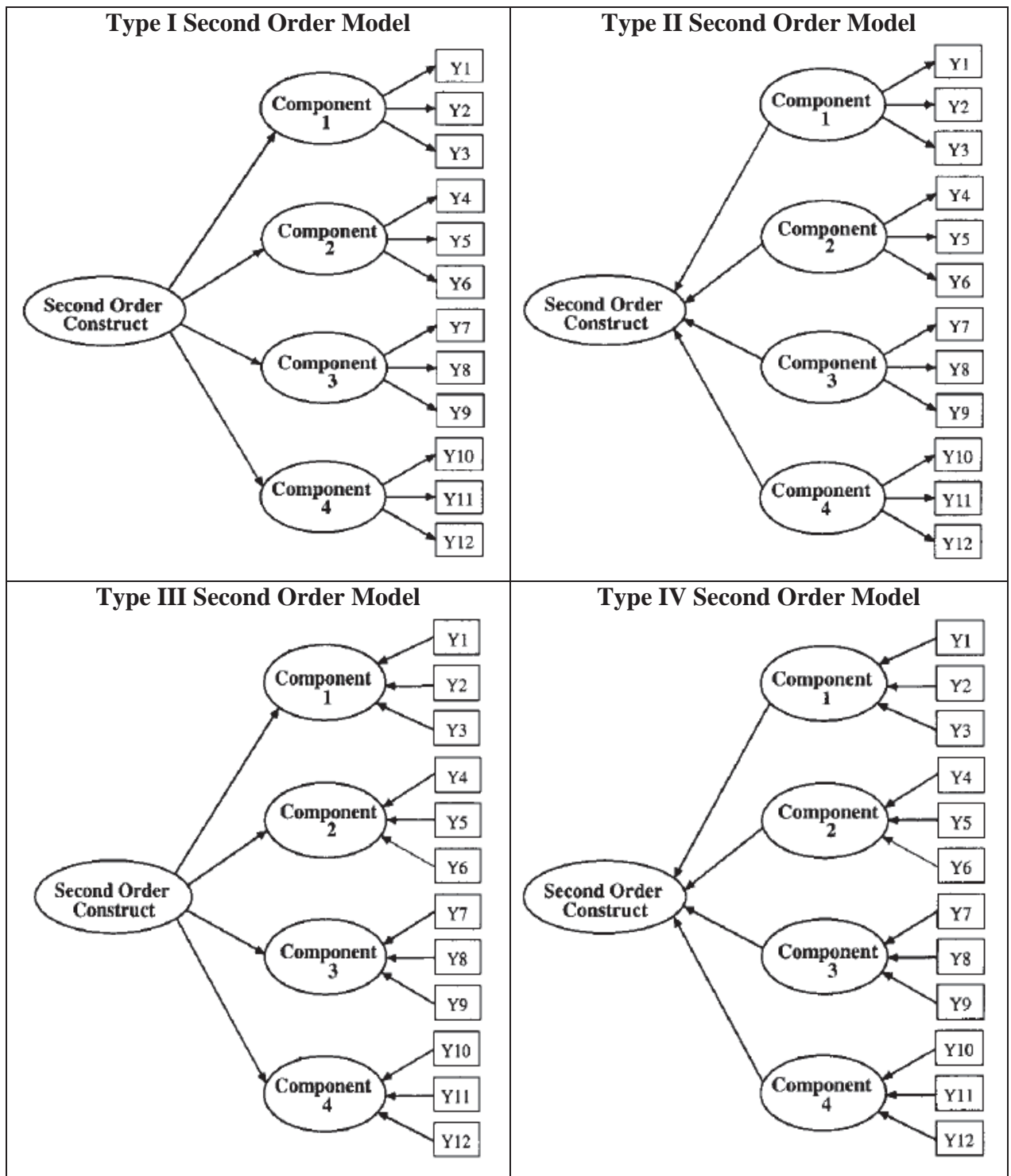


Figure 5.3: Alternative forms of representing a second-order construct (adopted from Jarvis, MacKenzie, & Podsakoff, 2003)

5.3. FORMS OF VALIDITY

Having shown how to represent constructs and research hypotheses both diagrammatically and statistically, in this section, the researcher discusses the concept ‘validity’ from a quantitative researcher’s standpoint, in relation to the measurement system, the constructs, the hypotheses, and the overall theory.

The ‘scientific utility’ of the measurement system is often referred to as the validity (or instrumental validity) of the measurement system (Nunnally & Bernstein, 1994, p. 83). Nunnally and Bernstein contend that validity assessment indicates how well the elements of the measurement system (e.g. the questionnaire items in a survey) measure what they are supposed to measure. There are three major types of measurement validity: content validity, predictive validity, and construct validity (Cooper & Schindler, 2014). Each type of measurement validity is reviewed in turn.

5.3.1. Content Validity

Content validity refers to the sufficiency and representativeness of the elements of a measurement system (e.g. the items in a survey questionnaire, the questions included in a test on a specific subject domain), given the domain of the subject that is being studied (Cooper & Schindler, 2014). The content validity of a measurement system is a matter of (subjective) judgement; content validity is usually judged by a panel of experts knowledgeable about the subject domain being surveyed (Haynes, Richard, & Kubany, 1995; Kline, 2011). Although content validity is not a statistical matter, statistical techniques are sometimes used to process expert judgements (e.g. see Lawshe, 1975; Wynd, Schmidt, & Schaefer, 2003). Face validity is also a form of content validity, but as the term implies, face validity takes a less stringent examination of the completeness of the elements (e.g. questionnaire items) included in a measurement system (Lynn, 1986; Nevo, 1985).

5.3.2. Predictive validity

Predictive validity, which is also referred to as *criterion-related* validity (Cooper & Schindler, 2014) or *empirical validity* (Nunnally & Bernstein, 1994), provides evidence of success of measures employed for prediction. Thus in predictive validity, the analyst attempts to predict an external criterion (sometimes referred to as a global measure) via

the measures used in operationalising the construct/s. Cooper and Schindle highlight four qualities that the chosen criterion should satisfy: relevance, unbiasedness, reliability, and availability. Since predictive validity (and even content validity) is subsumed in construct validity (see section 5.3.3 below), Straub, Boudreau, and Gefen (2004) assert that inclusion of a test on predictive validity in research reporting is optional. However, the researcher used a test on predictive validity to establish construct validity of a formative construct included in her model, because there are no other tests available to establish construct validity of a formative construct (Hair et al., 2014).

5.3.3. Construct Validity

Construct validity is defined as the extent to which the operationalisation/s of a construct/s measures what it is supposed to measure (Cooper & Schindler, 2014; Nunnally & Bernstein, 1994). Construct validity is being viewed by psychometricians as a very broad concept, which subsumes other forms of measurement validity, such as the predictive validity and the content validity (Kline, 2011; Mitchell, 1985).

Based on the literature (e.g. Bagozzi, Yi, & Phillips, 1991; Gefen & Straub, 2005), the researcher observes that *convergent validity* and *discriminant validity* are two important facets of construct validity. Convergent validity is confirmed when “each measurement item correlates strongly with its assumed theoretical construct” and discriminant validity is confirmed when “each measurement item correlates less strongly with all other constructs, except for the one to which it is theoretically associated” (Gefen & Straub, 2005, p.92).

5.3.3.1. Reliability and Internal Consistency

The researcher observes that the concept ‘reliability’ in validity testing of abstract concepts, evolved in the field of psychometrics in relation to ‘test scores’, in parallel with the definition of validity (Cronbach, 1951; Henson, 2001; Thompson, 2003). Whereas validity of the test scores means proof that measures that reflect an abstract concept measure what it purports to measure—based on the test scores—reliably of test scores refer to the extent to which a set of scores indicate the true scores of the abstract concept under observation (Henson, 2001; Thompson, 2003). Within this definition, it

is implied that reliability is a necessary, but not a sufficient condition for validity (Dawson, 1999; Thompson, 2003). The formal definition of reliability, which comes from classical test theory, is that reliability is the variance of true scores over the variance of the observed scores (Dawson, 1999). Supposing if X is the score of the variable under consideration and T is the true score of the variable for a given case, and E is the error term for that case (i.e. $X = T + E$), the reliability K can be defined as:

$$K = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{\text{Var}(T)}{\text{Var}(T) + \text{Var}(E)} \quad (5.5)$$

Since the true scores cannot be determined in practice, psychometricians have proposed several proxies for the above reliability coefficient K (see equation 5.5). The most well-known is the one developed by Cronbach (1951), known as the *internal consistency* reliability coefficient (also known as *coefficient α* or *Cronbach's α*). In essence, internal consistency is the extent to which measures of an underlying concept (construct) are internally consistent to yield similar results, for the concept under observation (Cronbach, 1951; Thompson, 2003).

5.3.3.2. Methods of Establishing Construct Validity

Factor analysis, either in exploratory factor analysis (EFA) mode or confirmatory factor analysis mode (CFA) is often being used to establish construct validity (Gefen & Straub, 2005). In the EFA mode, the researcher identifies the structure of the factors by examining which measures align (relate) more strongly to one factor compared to the other factor/s (Hair, Ringle, & Sarstedt, 2012). The numbers of factors that a researcher extracts from the total set of measures is based on the Eigenvalues of each extracted factor (Hayton, Allen, & Scarpello, 2004; Yong & Pearce, 2013). In the CFA mode, the factor structure is predefined by the researcher (i.e. the measures are pre allocated to the constructs/factors based on the operational definition of each construct) and the researcher's goal is to confirm this factor structure. The default technique for conducting CFA is SEM, which is discussed later (Gefen & Straub, 2005; Kline, 2011).

5.3.4. Statistical conclusion Validity

In the previous subsections (i.e. section 5.3.1 through to 5.3.3), the researcher reviewed the meaning of validity in relation to a measurement system involving latent variables (constructs). In this subsection, the researcher reviews the meaning of validity from the point of view of the hypothesised statistical relationships between latent variables. Cook, Campbell, and Donald (1979) defined the term ‘statistical conclusion validity’ to mean the validity of the statistical conclusions on the latent variable relationships (i.e. hypotheses) specified by the researcher. Austin, Boyle, and Lualhati (1998) assert that statistical conclusion validity is not only being governed by the true nature of the latent variable relationships that exist out there, but also by the statistical *power* of the test. The statistical power of a test is the probability of rejecting the null hypothesis in favour of a true alternative hypothesis (Lipsey, 2000). It is easy to show that the determinants of statistical power (in the context of the hypothesised relationships between the latent variables) are the size of the true relationship between the latent variables (the effect size), the sample size, and the significance level (Type I error) specified by the researcher (Cohen, 1992; Lipsey, 2000). Conversely, it is easy to calculate the sample size required to achieve a desired level of statistical power (typically $\geq 80\%$) (see Cohen, 1992 for a primer on the sample size).

5.3.5. Internal and External Validity

Thus far, the researcher reviewed the validity of a measurement system as well as the validity of the hypothesised relationships, purely from a statistical point of view. The researcher observes that the validity of a research design can be expanded over and above these domains. The two forms of validity in this regard are the internal validity and the external validity.

5.3.5.1. Internal Validity

Internal validity refers to the “extent to which conclusions can be drawn about the causal effects” of one construct on another (Hoyle et al., 2002, p.32). Therefore, internal validity of a research design concerns the ability of a research design to exclude rival explanations, such as the effects of factors not included in the design, on the observed phenomenon (Bryman, 2012; Malhotra & Grover, 1998; Rungtusanatham, Choi, Hollingworth, Wu, & Forza, 2003). It is documented in the literature (e.g.

Bryman, 2012; Cooper & Schindler, 2014; Malhotra & Grover, 1998; Straub, 1989) that establishing causality through nonexperimental methods such as survey research is a challenging proposition. Malhotra and Grover (1998) recommend that internal validity of survey research can be enhanced in two ways: *informally*, through a strong justification as to why the hypothesised relationships should exist (or why alternative explanations are implausible) and *formally*, by following certain specific procedures. As informal procedures, the following are proposed: the researcher's primary theoretical model is the causal model of the nonprofit balanced scorecard (BSC); the nonprofit BSC is a framework that has been subjected to intense academic scrutiny, although what the researcher examines is unique. In addition, the researcher provided the theoretical basis of her main model (Chapter Four) and conducted a case study to justify her constructs. The researcher also adopted two of the three formal procedures suggested by Malhotra and Grover to enhance internal validity: adopting a path analytic approach (the researcher used SEM) to test the model and observation of multicollinearity among variables. The third formal procedure of conducting a follow-up study (to examine causal effects longitudinally) was not pursued by the researcher due to time constraints, and this has been outlined as a limitation of the study. The researcher also observes that her research design guards against the following threats on internal validity, articulated in the seminal work of Campbell and Stanley (1963): *selection bias, attrition, history, maturation, and instrumentation*.

For these reasons, the researcher argues that action has been taken to enhance internal validity of the research findings.

5.3.5.2. External Validity

External validity refers to the extent to which the study findings can be applied or generalised to the populations and settings concerning the hypotheses (Hoyle et al., 2002). Since the researcher selected large representative samples, the study was shown to be free from method bias, and the findings were not influenced by the actions of the researcher (see Cooper & Schindler, 2014), the researcher argues that action has been taken to enhance the external validity of the research findings.

5.4. STRUCTURAL EQUATION MODELLING

Structural equation modelling (SEM) is an advanced second generation multivariate statistical technique widely used in social sciences to test theories and to validate the underlying operational definitions of the theoretical constructs (Bollen, 1989; Grigg & Jayamaha, 2014; Hair et al., 2014). SEM is a confirmatory approach, which provides a method for assessing and modifying theoretical models to confirm the hypothesised relationships between the constructs as well as relation between constructs and their measures (Anderson & Gerbing, 1988; Byrne, 2010; Raykov & Marcoulides, 2006). Unlike traditional multivariate procedures, SEM is capable of providing explicit estimates of the measurement errors of each measure (Byrne, 2010; Raykov & Marcoulides, 2006). An assumption that holds when using SEM is that the relationships between constructs (latent variables) as well as the relationships between constructs and their underlying measures (manifest variables) are linear (Raykov & Marcoulides, 2006). There are two types of SEM techniques: the covariance based SEM (CBSEM) and the partial least squares based SEM (PLSBSEM). Each technique has its own characteristics, assumptions, and therefore, advantages and disadvantages. These two SEM techniques are discussed in the next two sub-sections.

5.4.1. The Covariance Approach to SEM

The CBSEM—also known as Linear Structural RELations (LISREL), Covariance Structure Analysis, Covariance Structure Modelling, Causal Modelling, Analysis of Moment Structures (AMOS)—was invented by Karl Jöreskog (Jöreskog, 1970)³⁶ as a general method for analysis of the covariance structure of measures that operationalise the constructs. The purpose of CBSEM is to determine how well a researcher's hypothesised model is capable of replicating the covariance between the measures (manifest variables), through the estimated parameters of the specified model (see section 5.2.2.1). The CBSEM algorithm estimates model parameters through an iterative procedure, by minimising the discrepancy between the sample covariance matrix and implied covariance matrix (Bollen & Hoyle, 2012).

³⁶ LISREL and AMOS are also names of two popular SEM software packages available in the market.

The strength of the CBSEM technique over PLSBSEM technique is that being a ‘full information method’, the CBSEM algorithm estimates all the model parameters simultaneously, through a global optimisation parameter; therefore CBSEM provides an overall goodness-of-fit test for the specified model (Anderson & Gerbing, 1988). The main disadvantages of the CBSEM technique are that CBSEM requires: (a) large samples ($n > 200$), (b) mature operational definitions (not suitable for testing new/evolving operational definitions), and (c) measures to be multivariate normally distributed (Chen, Bollen, Paxton, Curran, & Kirby, 2001; Fornell & Bookstein, 1982). Chen et al. (1982) and others (e.g. Boomsma & Hoogland, 2001; Kline, 2011) have shown that violation of these requirements results in ‘improper solutions’ (e.g. non-convergence of iterations and infeasible solutions such as negative variances).

5.4.2. The Partial Least Squares Approach to SEM

The PLSBSEM, sometimes referred to as the *variance-based SEM approach*, was invented by Herman Wold (Wold, 1980), the mentor of Karl Jöreskog. According to Wold, PLSBSEM is a robust statistical technique for drawing statistical inferences when a researcher is unable to justify parametric assumptions of data (especially normally distributed data, which is not always achievable).

In PLSBSEM, all the model parameters are divided into *blocks* (subsets) and the parameters in one block are estimated using the parameter estimates of the other blocks, using the ordinary least squares regression technique. Consequently, the PLSBSEM technique minimises the residual error sum of squares of the endogenous constructs, on its way to estimating the relationships (structural regression coefficients) specified in the structural model. In this technique, the path coefficients and other unknown parameters such as the factor loadings are estimated iteratively, block by block, until variance reduction is no longer possible (Fornell & Bookstein, 1982).

Unlike the CBSEM technique, being a ‘limited-information estimation method’, the PLSBSEM technique does not have a well-accepted global goodness-of-model fit measure to determine to what extent a model fits to empirical data, overall (Anderson & Gerbing, 1988; Chin, 1998; Hair et al., 2014). Because of this, the quality of the overall model is judged through the evaluation of the structural model and the measurement

model via a set of nonparametric *evaluation criteria* (Hair et al., 2014). Table 5.1 depicts the systematic investigative procedure proposed by Hair et al, to evaluate the quality of the overall model in PLSBSEM applications. Chin (1998) also assert that it is a good practice to use multiple quality criteria to judge the quality of the overall model, given that there is no global fit measure in PLSBSEM.³⁷

As mentioned earlier (section 4.6.3 in Chapter Four), the PLSBSEM technique is more appropriate than the CBSEM technique when the phenomenon under examination is relatively new or the operational definitions of the underlying concepts are changing (Fornell, 1987; Hair et al., 2014; Wold, 1982). In addition, the PLSBSEM technique is also justified under the following conditions:

- (i) When the model constructs have a large number of measures (Bollen, 1989; Byrne, 2010; Chin & Newsted, 1999; Kline, 2011).
- (ii) When there is a theoretical need for formative constructs (Chin & Newsted, 1999; Diamantopoulos et al., 2008).
- (iii) When the necessary data requirements such as the sample size, multivariate normal distribution and independence of observations are not satisfied (Chin & Newsted, 1999; Hair et al., 2014).

Given the researcher’s objectives and the research design, it is evident that the PLSBSEM technique is more appropriate than the CBSEM technique for her study.

Table 5.1: A Systematic Model Evaluation Procedure for PLSBSEM (Source: Hair et al., 2014)

Evaluation of the Measurement Model	
Reflective Measurement Model	Formative Measurement Model
<ul style="list-style-type: none"> • Internal consistency (composite reliability) • Indicator reliability • Convergent validity (average variance extracted) • Discriminant validity 	<ul style="list-style-type: none"> • Convergent validity • Collinearity among indicators • Significance and relevance of outer weights

³⁷ Though Tenenhaus et al. (2004) have introduced a global goodness-of-model fit measure for PLSBSEM, it is not yet accepted by the researchers.

Evaluation of the Structural Model
<ul style="list-style-type: none">• Coefficient of determination (R^2)• Predictive relevance (Q^2)• Significance of path coefficient

5.5. VALIDITY CONSIDERATIONS IN CASE STUDY RESEARCH

The researcher used the case study approach initially to get a deeper understanding of the constructs of her main model; this enabled the researcher to operationalise the constructs. Therefore, the researcher reviewed the quality criteria prescribed by positivistic case study researchers—that is, researchers who advocate the use of the case study method to develop propositions and to understand the rich meanings of the underlying concepts embedded in the propositions. The researcher observed that positivistic case study researchers (e.g. Dubé & Paré, 2003; Voss et al., 2002; Yin, 2014) have redefined construct validity, internal validity, external validity and reliability in the context of case study research. The researcher observes that interpretive case study researchers on the other hand (e.g. Graneheim & Lundman, 2004; Guba, 1981; Lincoln, 1995) emphasise trustworthiness, credibility, confirmability, and data dependability to address the same requirements.

Construct validity in case study research relates to recognising correct operational measures for the theories being studied (Yin, 2014). According to Yin, construct validity can be strengthened by using three approaches: using multiple sources of evidence, establishing a chain of evidence, and having the draft case study report reviewed by key informants. In addition to the above three approaches, Voss et al. (2002) added a fourth approach: ‘triangulation’.

Internal Validity in case study research relates to searching to create a causal relation, by believing that some conditions lead to another condition/s, as distinguished from fake relationships (Yin, 2014). According to Yin, in case study research, internal validity can be increased using several techniques during the qualitative data analysis phase: matching the similar patterns, cross-checking the results to assure the logical order of the findings, and using diagrams to assist explanations (Riege, 2003).

External validity in case study research is the extent to which a study can be generalised. The generalisation of the study is not a statistical generalisation but an analytical generalisation, where the researcher tries to connect findings to previously established theory (Gibbert, Ruigrok, & Wicki, 2008; Riege, 2003).

Reliability in case study research is concerned with getting the same results with repetitive operations such as data collection procedures, for the purpose of minimising errors and biasness of a study. Reliability of a case study can be increased by using an interview protocol, recording interviews, and showing the interview transcripts to the respondents (Riege, 2003). In the case of the researcher's case study, the conceptual model and the qualitative questionnaire acted as the interview protocol.

5.6. CHAPTER CONCLUSION

In this chapter, the researcher reviewed positivistic techniques used in testing theoretical models, because these techniques feature heavily in the researcher's study, in achieving the research objectives. Statistical representation of an abstract concept was explained (section 5.2) and different facets of validity were reviewed (section 5.3). Since SEM is the method of choice in behavioural sciences to examine the two key components of a theoretical model—the measurement model and the structural model—the two prevailing SEM techniques were reviewed to examine which of the two SEM techniques is more suitable for the researcher's study. The researcher showed that the PLSBSEM technique is more suitable for her study than the CBSEM technique due to two major reasons (section 5.4). Firstly, the study involves new operationalisations (i.e. constructs whose meaning has been newly formed, using a case study). Secondly, the study involves a formative construct, which is difficult to represent through the CBSEM technique. Since the case study is an integral part of the researcher's study, the researcher reviewed how validity claims are made in case study research in a manner that does not conflict the validity claims used in positivistic research (section 5.5). In the next chapter (Chapter Six), the researcher presents and discusses the results of the case study.

CHAPTER SIX

QUALITATIVE DATA ANALYSIS AND RESULTS

Tis better to understand, than to be understood

A Prayer from St Frances

6.1. INTRODUCTION

In this chapter, the researcher presents and discusses the results of the case study, which is the qualitative part of the study. The qualitative study (Chapter Four outlines the case study design) was designed primarily to refine the researcher's *conceptual model* and operationalise the PM dimensions, for the purpose of theory testing through quantitative data collection.

The results presented in this chapter enable the researcher to posit which set of PM dimensions constitute an integrated PM system for NPOs (this relates to RQ1), what their operational definitions are (this relates to RQ2), and what theoretical relationships exist between the PM dimensions (this relates to RQ3). Therefore, the results and the accompanying discussion provided in this chapter form the foundation for providing more supportive answers to RQ1, RQ2 and RQ3 in a subsequent chapter (see Chapter Eight on test results).

This chapter is organised as follows. Section 6.2 provides data pertaining to the case study organisations, which are the units of analysis of the study. Section 6.3 provides qualitative data generated from the semi-structured questionnaire and the follow-up interviews; these data have been classified under particular themes and subthemes with the aid of the qualitative data coding software 'Interview Streamliner'. Section 6.4 elaborates and discusses these qualitative data, by way of providing quotes from the case study participants as appropriate, to add richness to the results. Section 6.5 presents the operational descriptions of the PM dimensions, explaining how the qualitative data lead to these descriptions. Section 6.6 explains how the operational descriptions led to the development of the quantitative questionnaire to collect

numerical data to test the researcher’s PM model as well as to test the validity of the PM system itself (PM system also includes the scoring system that the researcher developed via the ancillary model). Also, the contents covered in each questionnaire item are elaborated in this section. Section 6.7 covers the theoretical and practical implications of the case study findings. Finally, section 6.8 concludes the chapter, highlighting the key findings covered in the chapter.

6.2. SUMMARY DATA ON CASE STUDY ORGANISATIONS

This section shows the relevant data and information on the case study organisations and the participants who represented these organisations.

6.2.1. The Case study Organisations

Figure 6.1 depicts the specific healthcare subcategories to which the 9 case study organisations (each organisation was represented by one senior manager) belong. As mentioned in Chapter Four, 6 New Zealand organisations and 3 Australian organisations participated in the case study. Figure 6.1 indicates that all five subcategories belonging to nonprofit *healthcare* organisations are represented in the study, at least once.

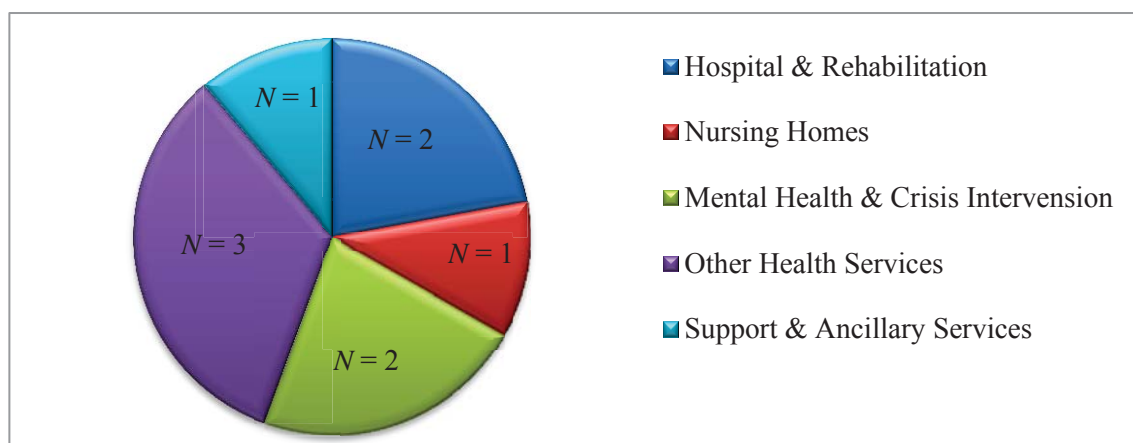


Figure 6.1: The case study organisations by subcategory

Figure 6.2 classifies the case study organisations by age (the age included the time the organisation has existed by other names, if any). This figure shows that only one organisation is new (less than five years old). All other eight organisations have been in

operation for over 20 years. Maturity in operation of the organisations meant that the respondents were able to provide very firm and authoritative responses to the semi-structured questionnaire and the follow-up interview.

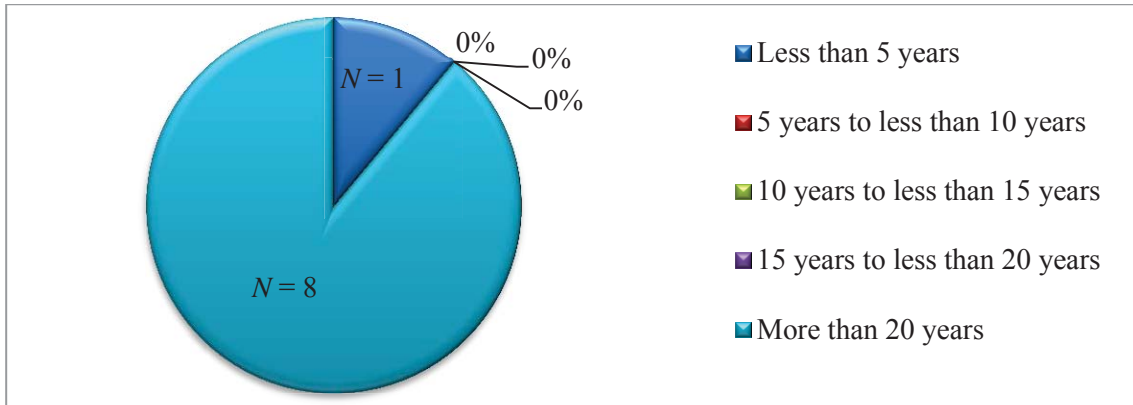


Figure 6.2: Classification of organisations by age

Table 6.1 depicts some useful information pertaining to the respondents and their organisations. All the nine respondents in the case study who represented their organisations were senior managers, because only senior managers were approached to participate in the case study. The participants were coded as follows:

- Senior managers belonging to the six New Zealand NPOs coded as N1 through to N6.
- Senior managers belonging to the three Australian NPOs coded as A1 through to A3.

Table 6.1: Profile of the Participants and Their Organisations

Respondent Code*	Organisational Category	Designation	Responsibilities	Familiarity with PM Systems**
N1	Hospital and Rehabilitation	Chief Executive Officer	<ul style="list-style-type: none"> • Manage the team dynamics • Managing the performance of fund-raising portfolios • Manage board • Ensuring relationship with DHB 	5 years
N2	Nursing Homes	General Manager	<ul style="list-style-type: none"> • Implementing strategies to sustain community based support • Implementing strategies to sustain residential and retirement village support 	10 years
N3	Mental Health	Senior Executive Officer	<ul style="list-style-type: none"> • Strategic development • Sourcing funding for continued and additional service delivery and training • Outcomes-based frameworks for service delivery 	10 years
N4	Mental Health	Chief Executive Officer	<ul style="list-style-type: none"> • To ensure the organisation fulfils the vision and meets the objectives of the strategic approach • Maintains a positive sector and public profile • Meets budget and remains sustainable • The organisation demonstrate the work and practices that are 	5 years

Respondent Code*	Organisational Category	Designation	Responsibilities	Familiarity with PM Systems**
			promoted: - Five ways to wellbeing - Mindfulness - Working Well	
N5	Other Health	General Manager	<ul style="list-style-type: none"> All Services 	4 years
***N6	Support and Ancillary Services	General Manager	<ul style="list-style-type: none"> Human Resources Learning and Research International Students Advocacy Communications and Marketing Quality Management Risk Mgt. Cultural Support Business Continuity Business Excellence 	10 years
A1	Other Health (Injury Care)	Chief Executive Officer	<ul style="list-style-type: none"> Guide and develop the organisation 	5 years
A2	Other Health (Anxiety Prevention)	Chief Executive Officer	<ul style="list-style-type: none"> Financial management Strategic management Human resource management Fundraising , corporate development and reporting to the board 	5 years
A3	Hospital & Rehabilitation	Director	<ul style="list-style-type: none"> Strategy, planning and governance 	10 years
<p>Notes: * A greater representation of New Zealand NPOs is due to logistical reasons (the researcher is based in New Zealand).</p> <p>** Familiarity of PM systems was questioned in the follow-up interview (the pilot testing of the quantitative questionnaire also provided this piece of information).</p> <p>*** N6 is a New Zealand National Quality Award winner</p>				

Information provided in Table 6.1 demonstrates that the respondents possessed wide-ranging experience related to strategy planning and strategy monitoring, which is essential in organisational-level performance management.

6.3. OUTCOMES OF THE SEMI-STRUCTURED QUESTIONNAIRE AND THE FOLLOW-UP INTERVIEWS

As mentioned in the methodology chapter, the researcher initially sorted the qualitative data (text) generated from the semi-structured questionnaire and follow-up interviews using the ‘Interview Streamliner’ software, which acts as a relational database system operating within Microsoft Access, which is part of MS Office Professional (Pruijt, 2012). The application of Interview Streamliner software begins by creating a text file by the analyst in *.txt format, using the specified syntax such as the ‘@’ sign and the underscore sign (_) described later. This enables Interview Streamliner to sort data under a particular keyword from the text file. The text file in this research contained written responses provided by the respondents and the interview transcripts. The ‘@’ sign should be used to separate ‘fragments’, while the underscore sign should be used to enclose keywords (Pruijt, 2012). In this study, a fragment represents the response provided by a particular respondent (the written response and the interview transcript), to a particular questionnaire item (question number). After importing the text file to Interview Streamliner, keywords can be recalled (i.e. searched) (Figure 6.3).

Since the questionnaire items were based on the conceptual model, data were identified as relating to one or more following keywords: *mission & core values, vision, strategy, economic sustainability, organisational capabilities and people development, processes, stakeholders, stakeholder expectations, stakeholder satisfaction, measures, contextual, existing PM systems, and miscellaneous*. The thematic grouping was the basis for the coding at this stage. The data were sorted for each keyword via the Interview Streamliner database. For example, after importing the text file containing fragments (the written responses and the follow-up interview responses that were transcribed), the researcher searched fragments under the keyword ‘Mission & Core Values’ and retrieved 13 records related to those keywords (Figure 6.3 depicts record 1 of 13). These were then exported to an external file, which was then used to search for patterns or similar ideas to group them together. This grouping was done manually by the researcher using highlighting pens, after obtaining the printouts of the external file on A4 paper.

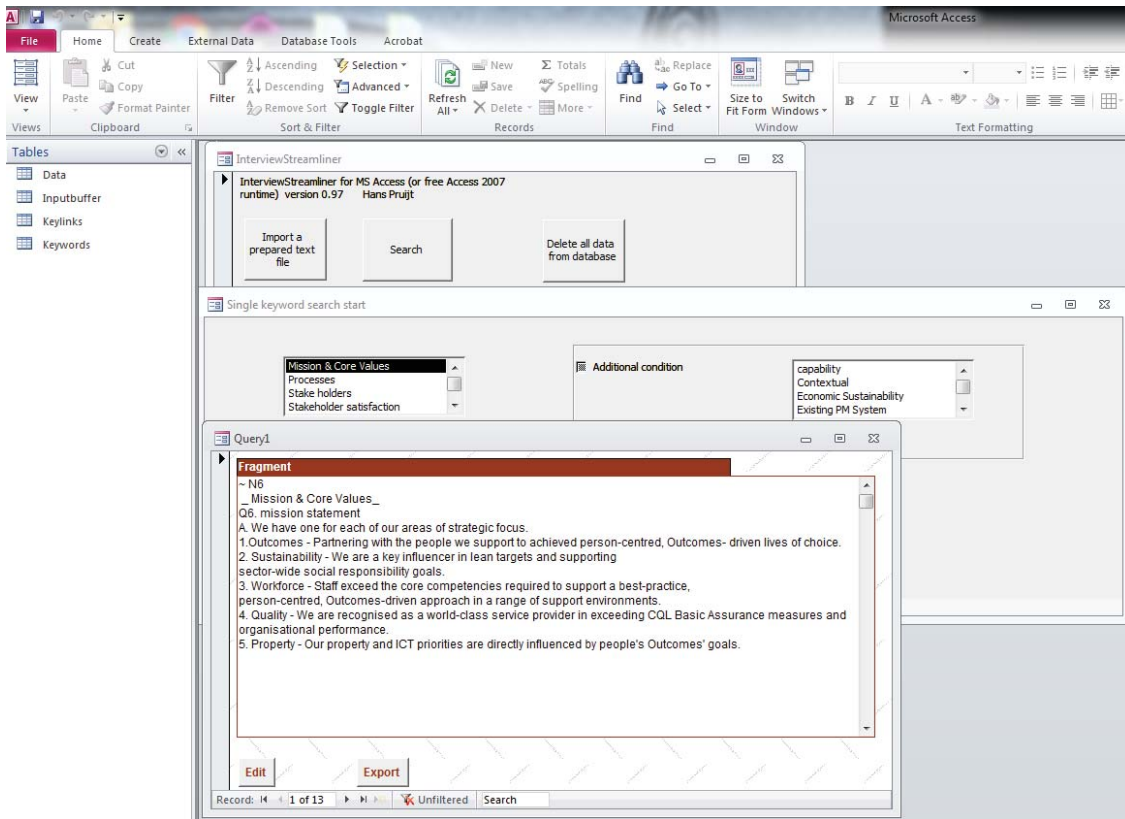


Figure 6.3: A screen dump of an output generated by Interview Streamliner software

Table 6.2: Questionnaire Numbers Corresponding to Keywords

Keyword	Questionnaire #	Keyword	Questionnaire #
mission and core values	6, 10, 21,11	stakeholder expectations	8
vision	10	stakeholder satisfaction	9,16, 21, 22
strategy	8, 9, 11, 12, 13, 16	measures	19, 20
economic sustainability	13, 14	contextual	5, 7, 9, 15, 16, 17, 18, 20, 21
organisational capabilities & people development	14, 15, 17,18, 19	existing PM systems	18, 20
processes	13, 14, 15	miscellaneous	1, 2, 3, 4
stakeholders	7		

Similarly, the researcher searched for all the keywords. As a result, the researcher received a specific set of qualitative data under each keyword for each case study

organisation. The *questionnaire number* corresponding to each keyword was also noted (Table 6.2). The researcher then scanned the text under each keyword to search for patterns or similar ideas to group them together. Each grouping was referred to as a subtheme, for the purpose of data analysis. Similar subthemes were grouped to form themes. These themes were then cross-referenced with the concepts (constructs) of the conceptual model, for the purpose of refining the conceptual model. The themes and subthemes relevant for the operationalisation constructs, along with phrases (the phrases came from the fragments) belonging to each subtheme are presented in Table 6.3.

Table 6.3: Themes, Subthemes and Related Information

Theme	Subtheme	Phrases Under Each Subtheme
Mission and Core Values	Social Profit	Supporting the community or a targeted group; provide social services that effectively meet the needs of those that the NPO serves; supporting people to lead fulfilling lives; contribution towards the overall wellbeing of the targeted group; trustable child healthcare; independence, dignity and unlimited opportunity for our targeted group; growing Kiwi kids; support for high needs community; improving the life of the clients
	Organisational and People Excellence	provider of choice; services of choice; world-class service provider; employer of choice; person-centred excellence; quality of service or the programmes; performance excellence; staff being effective in their work; maintaining or increasing clients' quality of life; organisation as the recognised leader in all matters relating to what we specialise in.
Strategy	Strategy Planning	The number of long term projects; developing a property portfolio; programmes proposed over programmes needed index (required actions); the number of new speciality services; the number of new strategies for fund raising; surveying the environment; developing new models of service provision; developing quality improvement initiatives; developing further resource allocation; developing and promoting organisation-wide leadership; developing client

Theme	Subtheme	Phrases Under Each Subtheme
		led services; developing long-term strategic objectives; creating a one-stop shop for advice; referral and services; staff involvement in SWOT analysis workshops (to see their concerns reflected in strategy).
	Strategy Implementation	Implemented projects; new staff recruitment and development programmes to fit strategy; further resource allocation to implement strategies; meetings before the final site business plan is been written and signed off; promotion of the service outcomes framework (the business plan is based on the service outcomes framework) in staff media; workshops on the strategic and operating plan for staff (this is to show how and where they fit); commend individuals and teams regularly through staff e-letters, Chief Executive updates, media articles and reports; cascading individual departmental operating plans to each employee to establish the link of their work to the organisation's strategic goals; a pliant agreement which has very clear service duties for the job which is important for what the organisation tries to achieve in the long term.
Organisational Capabilities and People Development	Board of Directors, Professionals and the Senior Leadership	High board and management alignment; governance and management experience; creating a (comparatively) highly qualified, high performing organisational culture; organisational management and professional, business-oriented decision making; a strong management team
	Training	Developing a high performing culture; producing trained (qualified) employees; training programmes and employee feedback on the effectiveness of such programmes; creating a committed team of highly engaged employees; ongoing training of evaluators; internal training; increased training to all staff; implementing a literacy programme to improve literacy levels and enable staff to acquire work-related qualifications; developing the organisation's training programmes.
	Infrastructure and Technology	Investment in the organisation's infrastructure; maintenance and investment in information

Theme	Subtheme	Phrases Under Each Subtheme
		technology (IT); review of the effectiveness of the IT; creating a fully integrated IT platform; establishing IT systems; developing a “paperless” culture; developing online recruitment processes; developing e-services.
	Marketing and Promotion	Brand promotion; investing in publications, advertisements, discussion papers, forums or symposiums; establishing shared projects, new partners and projects; creating brand awareness; promoting the organisation within the sector; creating (or discussing) a dialogue or a conversation in the media, and with the general public; growing knowledge and an interest in the public; reviewing organisational marketing material.
Economic Sustainability	Monitoring and Controlling Cash Outflows	Keeping a tab on the overhead costs; reducing property costs; monitoring and controlling average cost per paid and unpaid employee (volunteers); limiting total number of paid employees; monitoring the effectiveness of training costs; monitoring the cost for services, infrastructure and provision of IT.
	Revenue Increase	Establishing fund raising projects; monitoring the performance against budgeted targets; engaging in initiatives to create an interest in new contracts; maintaining good relationships with key individuals; being tactful in winning key funders (donors); constant engagement with funders; building relationships with the contractors, sponsors and donors.
Processes	Continuous Improvements	Continuous development of projects and services; creating work improvement teams (including external consultants when necessary) for efficient project management; generating publications, discussion papers and forums on continuous improvement in projects; implementing programmes for staff development to make continuous improvement more effective; ongoing training of evaluators; internal training on the service models; the “Baldrige Criteria for Performance Excellence” approach being embedded in organisational policy; implementing

Theme	Subtheme	Phrases Under Each Subtheme
		alternative philosophies.
	Designing of Safe, Efficient and Effective Processes	Ongoing review of costs and revenues pertaining to process re-design (i.e. examining their effectiveness); cost saving but effective processes; bulk purchasing; developing and implementing further resource allocations for process improvement; increasing workforce capability; establishing a robust reward and recognition system; strengthening or improving the quality of the processes (“strengthening” means embedding quality at the process design stage while “improving” means quality control/checks in continuous improvement); implementing staff health programmes; on the job training for high impact and complex support staff; developing mentor training for service manager and service coordinator roles; reviewing complaints processes; reviewing client outcomes tools; workforce development.
	Designing the Infrastructure, Technology and Materials	Establishing well developed IT services; developing online recruitment; Investing in IT infrastructure; equipment purchases over equipment request index; reviewing services that could not performed due to lack of resources; information systems innovation; standardising information systems; developing a IT steering group; reviewing marketing and promotion material for effectiveness.
Stakeholder Satisfaction	High Quality of Support and Services	Services offering feedback, and family surveys; establishing internal measures on achievement and growth; conducting internal and external service quality audits; consultations/treatments (measuring quality of life); investing in people (paid employees); investing in recruiting volunteers; developing training programmes; increasing the skills set (qualifications) of paid staff; achieving/targeting a staff satisfaction rate of 85%; achieving/targeting a client satisfaction rate of 90%; remaining the preferred provider of service at the high end of disability; measuring subjective quality of life; client outcomes assessment and tracking the satisfaction rates

Theme	Subtheme	Phrases Under Each Subtheme
		<p>over time; conducting service reviews with clients and their families; deploying assessors to assess client’s goals and outcomes; commissioning clinical indicator reports; conducting social return on investment research and other relevant research on measuring social impact.</p>
	<p>Skilled and Recognised People</p>	<p>Establishing reward and recognition schemes; conducting performance appraisals; conducting workshops for staff educating them on the strategic and business plan; commending the contributions of the employees (volunteers included) through media and internal articles and reports; conducting performance reviews (appraisals) on paid employees annually; evolving a skilled and competent workforce; providing feedback to employees on their progress against the objectives set by their line manager.</p>
	<p>Well Performing, Empathic Social Service</p>	<p>Conducting audits and surveys to check the organisational climate; collaborating with other organisations; establishing additional programmes such as counselling, parenting, etc.; doing “extra miles” in service provision; relying on the turnover index; attending to meetings with donors and policy holders; developing and implementing client-led services; developing and implementing new models of service provision; partnering with others to help them on their performance excellence journey; basing the vision of the organisation and its strategic goals entirely upon social obligations; fulfilling the social responsibility, by definition.</p>

6.4. DISCUSSION OF KEY THEMES AND SUBTHEMES

In this section, the researcher discusses the key themes identified through qualitative data analysis (Table 6.3), providing quotes from the respondents as much as possible, to add to the richness of qualitative data.

6.4.1. Mission and Core Values

The qualitative data (Table 6.3) confirms that Mission is at the heart of NPOs and that NPOs are mission driven organisations. For example N6 mentioned the following:

We are mission driven organisations and value driven too. Our mission, vision and values are driven from the top. Our recruitment process, reward and recognition systems and our operational processes revolve around our core values. Everything is aligned to achieve the mission.

All participants explicitly stated that the mission statement of their organisation connote that they are a *social profit* organisation, rather than a passive entity. For example, N1 stated “*everybody talks about not-for-profit but for me it’s about social profit. We exist to generate money to put back into society, to make benefits to the society*”; in a similar vein, N5 stated that “*our bottom line isn’t profit; our bottom line is the intangible outcome for the people that we support*”. The researcher observes that this social profit notion is in line with the contemporary literature (e.g. see Maas & Liket, 2011). The participants also conveyed that the basis of their mission is the community/target group requirements. All the participants mentioned that their skilled staff are an organisational strength and that staff (and volunteer) development is embedded in their mission. In addition, employees’ passion for mission, quality of services, the variety of services, the efficiency and effectiveness of organisational performance emerged as core values. The researcher also observes that, because core values are invariably part of the mission, the measurement domain Mission and Core Values can be relabelled as Mission.

All participants explicitly stated that they exist to provide benefits to the society or their targeted group. Respondents used vivid phrases—the “*provider of the choice*” (N4, N6 and A1), the “*leader*” in their particular service area (A2, A3 and N3), an organisation that pursues “*sustained excellence*” in healthcare (N6)—to highlight the social dimension embedded in the mission of their organisation. The researcher observes that the “sustained excellence” notion (maximising stakeholder value) in a nonprofit context stands analogous to sustaining a steady stream of net cash inflows to maximise shareholder value in a for-profit context, as posited in the “theory of the firm” (Brigham & Pappas, 1976; McWilliams & Siegel, 2001). All the participants mentioned that the “high-quality of service” notion is embedded in their mission statement and that they need skilled teams, consisting of people belonging to all levels of their organisation. This means, high performing organisations advance their staff to create and deliver efficient and effective services which result in satisfied customers. Thus, the qualitative study reveals that *Social Profit* and *Organisational and People Excellence* capture the conceptual domain of the construct Mission. Therefore, the researcher advocates that NPOs need to work in these two subdomains to fulfil their mission.

6.4.2. Strategy

Strategy is a critical factor that influences organisational performance. All participants explicitly stated that, since goals and key projects originate from the strategy and since the mission and vision are embedded in the strategy, it is important to communicate the strategy effectively across the organisation, and link departmental and individual goals to strategy. The respondents also emphasised the importance of not only the ‘top-down’ but also ‘bottom-up’ sharing of information regarding strategy. Further, they mentioned that it is also important to make their people (employees and volunteers) aware that their efforts contribute to achieving the objectives of the organisation. Some extracts from participants’ responses are as follows: “*all staffs are involved in SWOT analysis workshops and they see their concerns reflected in our strategy*” (N6); “*during orientation, staff participate in workshops on the strategic and business plan; so they become more aware of these and how and where they fit*” (N5); “*the individual departmental operating plans are cascaded to each employee, so that they can see the link of their work to the organisation's strategic goals*” (A1); “*we are transparent with*

all employees and volunteers as to the operation of the business, and we involve them in all parts of the business; they are part of the goal setting process” (A2).

The participants accepted that monitoring the currency of the elements of the Strategy itself is part of developing a PM system. Some extracts from participants’ quotes in this regard are: *“it is fundamental to our performance” (N1); “strategy is everything, and as the saying goes, if you don’t know where you’re going, you’ll end up somewhere” (A1); “we place great emphasis on getting our strategy right, and making sure that all employees’ actions are focused on delivering the strategy” (A3); “the strategy is focused on increased performance and has specific measures that relate to those, for instance” (N2).*

In general, these statements imply that *Strategy Planning* aspects as well as *Strategy Implementation* aspects should be covered under the PM dimension ‘Strategy’. Coding of follow-up interviews and written responses identified the following as important PM elements on strategy implementation: communication or the communication plan; action plans for the objectives and how they align throughout the organisation; references or the guidelines used and planning decisions, business drivers and how they were derived; and features designed to facilitate the implementation of the strategy.

6.4.3. Organisational Capabilities, and People Development

To obtain a richer perspective of the Organisational Capabilities and People Development dimension, the researcher included a series of questions in the semi-structured questionnaire (Q14 through to Q19) instead of asking the single question: “what capabilities do you develop to drive key processes to deliver stakeholder value?”, a key question related to the Prism (Neely, Adams, & Kennerley, 2002). However, in the follow-up interviews, the researcher did ask this question in order to cross-validate the responses. The most frequently cited resources were *the Board of Directors (n = 8); Professionals and/or the Senior Leadership (n = 9); Training (n = 6); Infrastructure and Technology (n = 5); and, Marketing and/or Promotion (n = 4)*. The respondents N2 and N5 used the phrase *“creating brand awareness”* to describe *Marketing and Promotion* more precisely. The respondents also mentioned that their organisation’s information infrastructure acts as the medium through which tangible and intangible

resources drive the processes. The key quotes from participants' responses on resources are mentioned below.

Our staff do a good job; we highly value their contribution; coz we have an internal training system and we identify most of the training needs from our internal audits; the audits identify things that we ought to put into training to tap human resource development parameters. I'd say that we have a comprehensive management information system that identifies human resource development needs (N2).

There are number of factors that affect performance. One is the shortage of skilled staff. I mean people who are qualified and have the required knowledge around managing disability services. This is a big challenge to us. The second factor is a good IT system. We have got quite a sophisticated IT platform for reporting our achievements towards the mission, performance, incidents, and tracking people. We provide training for legacy staff to upskill using our IT platforms. (N6).

We do client surveys to gauge our brand awareness, because we want to make sure that the donating clients are happy with what we have been doing. We ask donating clients what they know about us and what service we provide. Because we are in (the) healthcare, there is a perception among donors that we get heaps of funding from the government. The truth is that that we get none. Therefore, we have to increase our brand value through marketing and promotion. (N1)

We need to make sure that our teams keep going and our IT system keeps going, helping information sharing to improve our performance. (A2)

Five participants (N2, N3, N6, A2 and A3) mentioned that IT is a key elementary physical infrastructure that enhances organisational capabilities. As such, the researcher re-labelled the PM dimension under review as Organisational Capabilities, Infrastructure, and People Development to reflect the role that IT plays as a key physical infrastructure. From the pattern of responses on Organisational Capability, Infrastructure, and People Development, the researcher observes that the said PM dimension can be represented as two related dimensions: *People* and *Organisational Infrastructure*. This also requires an adjustment in the corresponding tentative causal links shown in the conceptual model. The researcher also observes that People and Organisational Infrastructure dimensions represent the Human Capital and Information Capital subdomains belonging to the Learning and Growth dimension of the BSC (e.g. see Kaplan & Norton, 2004). The extant literature (e.g. Akingbola, 2013; EFQM, 2013; Gurd & Gao, 2007) can also be used to justify distinguishing People from the physical resources and IT. Mission achievement heavily relies on the capabilities and the effort of the people (paid employees, professionals, and volunteers), because of the labour intensive nature (as opposed to capital/technology intensive nature) of the activities that are being performed in NPOs.

For Q17, all respondents cited their people as the major strength and/or the major capability of their organisation; in the follow-up interviews, they also mentioned the importance of staff training, which also includes training volunteers (the verbal responses that emerged from Q17 were consistent with the written responses that they provided for Q14). Some of the important statements found in the written responses include: *“People! People! People!: the staff we have, the people we contract with, the board, friends and partners of our organisation”* (N4); *“the staff we employ—nurses, health workers, others and our volunteers—they all have a passion over the organisation”* (N5); *“a strong management team, a committed team of highly engaged employees”* (A1); *“the major strength of the organisation is amazing clinical work staff”* (N1); *“our human resource planning department brings the right people in and our learning and development department lifts the skills of them”* (N6).

The follow-up interviews revealed that a NPO has to be led by its directors and the senior leadership for better performance (keeping mission/vision achievement as the end goal that they pursue) and that understanding its people is a major requirement in a

healthcare NPO. This highlights the need of development of the human resources of the NPO to enhance Organisational Learning and Growth. The participants also mentioned that human resource development creates a high performing work culture. These ideas encompass the construct 'People'. There is abundance of literature to argue that individual capabilities as well as team capabilities can be enhanced through formal training, tacit learning/knowledge and managerial innovation for service sector effectiveness (Erden, Von Krogh, & Nonaka, 2008; Holste & Fields, 2010; Hume, Clarke, & Hume, 2012; Shekar, 2007). The reader will note from Table 6.3 that the qualitative data collected via the case study under the subtheme 'Training', represents these facets of capability referred to in the literature. The researcher also found through follow-up interviews that people in NPOs have relatively low tolerance for work ambiguities and that they are motivated when they know that they can fall back on their leaders. Similarly, the qualitative data revealed that people are motivated when they perceive that high-level (strategic) decision making is 'outcome oriented'.

Through follow-up interviews, the researcher found that the construct Organisational Infrastructure acts as the medium that facilitates Organisational Learning and Growth and that Organisational Infrastructure encompasses *physical infrastructure, information and communications technology (ICT)*, as well as *marketing and promotion*. By investing in marketing and promotion of their brand name, organisations are expected to increase the recognition of the services that they provide to society, which attracts more donations/investments. The researcher's findings also agree with the definition of the 'organisational capability' given by the Neely et al. (2002). Neely et al. define organisational capability as a "combination of an organisation's people, practices, technology and infrastructure that collectively represent that organisation's ability to create value for its stakeholders through a distinct part of its operations" (p. 177).

6.4.4. Economic Sustainability

Although NPOs are not profit oriented organisations, they require financial capital to continue to serve their customers, pay employees (non-volunteers) and meet numerous overheads—in short, maintaining their organisations. This was the basis for including the dimension 'economic sustainability' in the conceptual model. The researcher observes from her qualitative data that the concept economic sustainability is embodied

in the *social profit* notion in the mission; this also implies that there is a direct link between Mission and Economic Sustainability, thus requiring further adjustments to the researcher's conceptual model. As Kaplan (2001) asserts, the financial sustainability of a NPO acts as an enabler for achieving the mission (through designing appropriate processes); for example, to create the requisite impact on the community, satisfy/motivate people in the organisation, and satisfy the donors (more generally, to satisfy the key stakeholders).

The written responses to Q13 and Q14 and the subsequent follow-up interviews provided evidence to justify the importance of a healthy financial infrastructure for a well performing organisation. Two important aspects emerged from qualitative coding. The first aspect was about processes directed at *Monitoring and Controlling Cash Outflows (cost control)*. The second aspect was about the initiatives taken by the organisation to *Increase the Revenue (cash inflow)*. The latter includes fund raising projects, donations, bequests and related activities. Further, increasing revenue can be split into two processes: processes directed at *increasing cash inflows* and investing in the right tangible and intangible resources. To align with the two resources People and Organisational Infrastructure embodied in the Organisational Capability, Infrastructure, and People Development dimension, the researcher relabelled the Economic Sustainability dimension as 'Financial Health'.

The researcher also found that it is important to maintain strong relationships with key stakeholders who provide cash inflows: funders, donors, and sponsors. These relationships include providing them with information and feedback on how funding was sourced to achieve the mission and objectives of the organisation. Some quotes from participants' responses were: "*we need to maintain high donor retention levels to sustain funding to serve our patients; we have specific donor retention targets specified in our financial objectives*" (N1); "*we build and maintain relationships with our contractors, sponsors and donors – having open transparent dialogue and providing them with feedback on how their funds have been used*" (N5). Three respondents (N2, N6, and A1) whose organisations heavily depend on external funds (government and private donations) mentioned that due to external pressure, they have to work to meet the agendas of external funders. This is also highlighted in scholarly literature (Froelich, 1999).

The qualitative data also revealed that a NPO typically increases its cash inflow through fundraising projects and continuous innovation to create new revenue streams, while it typically controls cash outflows by reducing waste and termination of activities that do not result in serving the customer/client. In addition, the researcher found that the ‘attractiveness of the mission’ is a factor that contributes to Financial Health, through increased cash flows. For this reason, the researcher introduced a direct effect from Mission to Financial Health in finalising her model (Figure 6.4).

6.4.5. Processes

The qualitative data generated from survey questions Q14 (the primary question), Q13 and Q15 surfaced three important process facets: *Continuous Improvement*; *Design of Safe, Efficient & Effective Processes*; and *Designing the Infrastructure, Technology and Materials*—the first two facets reflecting the incremental approach while the latter reflecting the big-step quality improvement/innovation approach referred to in the quality literature (Ciasullo & Troisi, 2013; Grigg & Mann, 2008).

Continuous improvement was found to capture the ongoing activity that develops over time, in manifesting a better performing organisational culture; design of safe, efficient and effective processes was found to capture the collaboration processes that encourage the use and sharing of best practices for better performance; and, designing the infrastructure, technology and material was found to capture the extent to which the human resource, finance, and ICT systems have been designed to support continuous improvement (and where relevant big-step) processes. Specifically, safety is crucial in the healthcare context because faults may lead to severe consequences. It is also highlighted the organisational improvements through collaboration processes that encourage the use and sharing of best practices which ultimately improve the entire sector growth. It is also important to design and invest in infrastructure, technology and material, which is important to provide a high quality service to the customers/clients and fulfil social responsibility. Achieving excellence in healthcare was found to be part of the mission in case study organisations. Furthermore, it is clear that effective and efficient processes cannot perform without the support of a suitable infrastructure and supporting technology and material. Thus the construct Processes was modelled to

capture the facets of Internal Business Processes, the label given in the BSC. Some important quotes which support these facts are shown below.

We establish partnerships with other nonprofits within our sector to strengthen ourselves as well as our sector. So we are big in sharing resources. We provide free training to people belonging to our partners to upskill them. So in the short term, we compromise our competitive advantage in trying to lift our partners. In the long term, we all benefit as the sector grows. (N4)

Without efficient and effective processes and resource capabilities, a strategy cannot be delivered, and the performance suffers. We use our processes to develop the capability to deliver on the strategy, and without efficient processes, an organisation cannot perform at the desired level. (A1)

We collaborate with other organisations in our sector, in certain projects or, help them in kind towards their projects. That's how we work. We would like to work together, but when it comes to fundraising, we remain competitive! (N1)

We put systems and processes in place to ensure that we have the right people doing right things at the right time in right numbers. We provide the right climate for our people to perform optimally; we ensure that we have enough people to work in our projects. We conduct performance reviews to identify people development needs and we have reward systems to motivate people. So we get the best people, the best skills and the best attitudes, combining into best processes. We also have several other operational strategies to improve our processes, but these are the key ones. (N6)

6.4.6. Stakeholder Satisfaction

Stakeholder satisfaction pertains to how well the services fit into stakeholder expectations. The coding of responses of the relevant questions that capture Stakeholder Satisfaction (Q9, Q21 and Q22 were the main questions but several other questions also provided information on Stakeholder Satisfaction indirectly) confirmed that Clients, People (employees and volunteers), and the Donors are the three major stakeholder groups the NPOs need to satisfy. This is consistent with the conceptual model. All respondents mentioned that a high quality of support/service delivery is of paramount importance for the success of their organisations and that quality of service is more important in healthcare than in other nonprofit endeavours. The researcher found that clients expect a safe, responsive and an exemplary service. One PM practical problem in service delivery that was uncovered in the follow-up interviews was that because the actual clients of the organisations are—in many instances—patients who suffer from various disabilities (e.g. frailty, mental health problems), the clients themselves are not in a position to express their satisfaction or otherwise with the services provided by the organisation. This means that organisations have to rely on feedback provided by the client's family or their support person to gauge the effectiveness of their services. Some quotes that justify the researcher's findings on client satisfaction are given below.

The families of patients that we look after want to know whether their members are being well looked after. Our patients are old and frail. They are unable to communicate with us, and therefore, by default, their families become our clients. These families are important stakeholders. Our goal is to improve our clinical performance to minimise patient illnesses or rational harshness. We conduct several internal and external audits to examine the effectiveness of our clinical processes. I think, we have a very robust quality system. We maintain a central register that monitors the progress of key activities on a monthly basis. The register also records people responsible for each activity. (N2)

We conduct surveys involving the families of the people that we support. We interview them to find out whether or not they are happy with the support we provide to our clients. Often, we need to deal with people who can communicate but not very effectively, because they've got either physical disabilities or intellectual disabilities. We use trained staff to communicate with our clients to find out what they feel about the support services that we provide. The feedback collected by my staff comes direct to me or to the board. Sometimes we interview our staff to check whether they have tweaked the feedback to please us! We want honest feedback. (N6)

We directly measure subjective quality of life through quarterly client outcomes assessments, and track this over time. This is a quality of life tool developed in-house. The tool measures the overall quality of life, as well as quality of life by each of eight life-domains, involving 41 measurement items in total. The findings of these assessments are reported to each level of the organisation, but the content that is reported to each level varies because information needs vary. For example, what a support worker needs to know will be different to what the board of directors want to know. (N3)

The researcher also found that the paid employees expect fair remuneration and growth opportunities; paid employees and volunteers both expect a flexible, supportive and caring organisation. They like to work for organisations that are willing to satisfy their people. They want to work in organisations that value the contribution of their work. The researcher admits that the case study did not cover interviews with staff. *People Recognition* is an important aspect that emerged from the study (qualitative data). Identifying high performing employees and rewarding them (both intrinsically and extrinsically) for their performance creates a strong bond between the organisation and the employee, which leads to increased levels of employee retention. The two quotes from the respondents that support this proposition are given below:

We often aim for 85% staff satisfaction and 90% IT user satisfaction; that reflects internal customer satisfaction. These are two of our KPIs.

We believe that a satisfied worker is a motivated worker who performs better, but also stays with the organisation longer. By losing a key person, we lose the unique skill set that person brought to the table. (N2)

We survey our staff to find out what they think about their own performance through reflection, and the effectiveness of the training given to them, and the benefits that they have been getting. We also survey the effectiveness of the senior leadership in fulfilling the needs and wants of the staff. (N6)

In regard to donors, the researcher found that donors expect a well performing, empathic social service. In addition, the researcher found that NPOs achieve donor satisfaction by designing their services around the needs of the community (or targeted group), developing and implementing new service models, and collaborating with other organisations to fulfil social responsibility. Key quotes on the donor perspective were: *“donors give us money for a specific purpose” (N2); “people come and would say here’s my donation and I want it to go specifically to the oncology ward because my son or the daughter or the nephew or the niece or the grandchild was treated at your hospital for leukaemia or cancer and I support that cause” (N1); “financial donors want to know whether donations are being used to purchase critical equipment or establish services that will improve the healthcare that we provide” (A2).* The participants also conveyed the importance of providing performance information to their donors to convince them how their contributions were used to achieve their mission. As in the case of people, the researcher acknowledges that donors were not interviewed by the researcher.

Due to different expectations of the three stakeholder groups, Client Satisfaction, People Satisfaction, and Donor Satisfaction were represented as three separate constructs. It is important to note that managing different and often conflicting stakes of multiple stakeholders is inherently challenging (Elias, 2015; Knox & Gruar, 2007; Krashinsky, 1997). For examples of managing a complex Australasian stakeholder in a nonprofit context (R&D) see Elias (2015).

6.5. PRESENTING THE OPERATIONAL DESCRIPTIONS OF THE UNDERLYING PM DIMENSIONS

Based on the analysis of qualitative data (Table 6.3) as well as the results that were discussed in the previous section), the researcher provides the operational descriptions of each concept (PM dimension in her model) as follows (Table 6.4).

Table 6.4: The Operational Descriptions of Each PM Dimension

Performance Measurement Dimension	Operational Description
Mission	1. A NPO exists to achieve social outcomes; social profit should forefront
	2. Improving the experience of the community should be highlighted in the mission
	3. Staff development should be highlighted in the mission
	4. Need to pursue leadership in what a NPO does to energise staff
Strategy	1. Triple bottom line elements—donors, community, and the environment—should be considered in strategy planning
	2. Senior leadership involvement is important not only for strategy planning but also for strategy implementation
	3. Achieving social responsibility requires collaboration with key stakeholders, pursuing continuous improvement
	4. Organisational structure and resources are considered in strategy implementation
	5. In operational planning, work has to be mapped to people to motivate them to perform
People	1. Senior leadership’s capability gives strength to the organisation to drive its processes
	2. Decision making should be “business like”; “nonprofit” means non-profit initiatives
	3. Creating a high performance work culture is as important to a NPO as for a for-profit organisation

Performance Measurement Dimension	Operational Description
Organisational Infrastructure	1. Information technology (IT) is vital for developing new services and communicating with the stakeholders
	2. Creating “brand awareness” and public awareness via the information system is a critical organisational capability
Financial Health	1. Investing in IT/infrastructure strengthens the financial position as this improves productivity, which in turn attracts funding
	2. Cost reduction (direct costs and overheads) improves the financial position
	3. Providing regular feedback to sources who provide funding increases the chance of sustaining the funding
	4. Relationships with stakeholders who provide funding is important for financial sustainability
	5. Innovation initiative development of new services attracts of new funding
	6. Investing in IT/infrastructure strengthens the financial position as this improves productivity, which in turn attracts funding
Processes	1. Process improvement should be aimed at achieving better outcomes for donors; utilising infrastructure, technology and materials
	2. Process improvement should be aimed at achieving better outcomes for clients; safe work processes are part of this
	3. Operational strategies are as important as corporate level strategies
	4. Support processes are part of continuous improvement
	5. New product development should take into account the community and the funder requirements
	6. Processes should also include IT/technology development
	7. Providing the resources and creating the right organisational structure is vital in strategy implementation
	8. The effectiveness of strategies need to be monitored and ineffective strategies need to be replaced

Performance Measurement Dimension	Operational Description
Client Satisfaction	1. Intense service audits usually result in client satisfaction; sometimes donor satisfaction also follows
	2. Quality of life of clients' needs to be measured/audited to assess and improve their satisfaction
	3. Association to Client feedback on the services results in their satisfaction
	4. Client satisfaction needs to be measured regularly
People Satisfaction	1. Paid employees should be remunerated appropriately
	2. Training improves employees' satisfaction
	3. Volunteer development programmes must not be overlooked
	4. Communicating strategic business plans to staff motivate them intrinsically
	5. Staff satisfaction need to be measured regularly
	6. Recognition of employees and volunteers improves their motivation and satisfaction
Donor Satisfaction	1. Processes designed around client needs usually result in client satisfaction, which ultimately results in donor satisfaction
	2. Developing new service models results in donor satisfaction (because this usually results in client satisfaction)
	3. Fulfilment of social responsibility requires collaboration with other organisations
<p>Note: The fifth operational description of the Financial Health (“fund raising projects”) was isolated as a global measure of Financial Health. In the quantitative questionnaire this operational description corresponds to question #4 (Q4). The technical reason for having to have a global measure for Financial Health is given in Chapter Eight (see section 8.2.4.1).</p>	

The 40 items listed in Table 6.4 (Q4 was excluded because it was identified as a global measure of Financial Health) along with the descriptive information provided by the researcher on each PM dimension (section 6.4) can be treated as the operational

definitions of each PM dimension (construct) in the researcher’s theoretical model. This answers the second research question (RQ2).

The themes (concepts) that the researcher identified were reconciled with the conceptual model to refine and finalise it. This included modifying the labels of the constructs (PM dimensions) of the conceptual model where appropriate, and adding new relationships between the constructs. The researcher’s final theoretical model is shown in Figure 6.4. Due to the availability of the operational definitions of the constructs and specified theoretical relationships, the model becomes testable for the validity, through a scientifically developed measurement instrument (section 6.6). In developing the measurement instrument, the researcher achieves her second research objective.

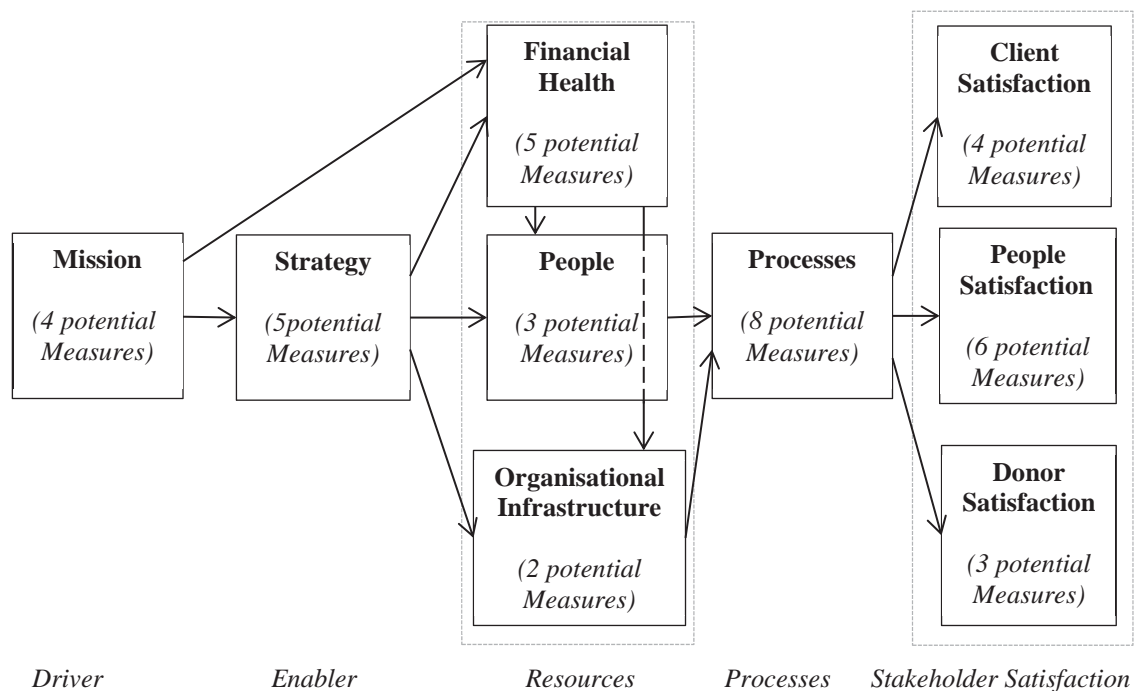


Figure 6.4: The final theoretical model

In a subsequent section (Table 6.6 in section 6.7), the researcher suggests suitable key performance indicators (KPIs) for each operational description shown in Table 6.4. These KPIs collected through the case study (the qualitative questionnaire) can be treated as one of many practitioner-related outcomes of the study.

6.6. DEVELOPMENT OF THE SURVEY ITEMS OF THE QUANTITATIVE QUESTIONNAIRE/MEASUREMENT INSTRUMENT

Having developed the 40 operational descriptions of each PM dimension/construct (Table 6.4), in this section, the researcher explains how these were translated to a form that suits a quantitative questionnaire to collect numerical data. Basically, each operational description summarised in Table 6.4 was translated into a format that can be clearly understood by a senior-level manager. As a specimen translation, the nexus between the first operational description of the first PM dimension (Mission) and the corresponding statement in the survey questionnaire is shown in Table 6.5.

Table 6.5: A Specimen Translation to the Questionnaire Format

Operational Description (in summary format as per Table 6.3)	The Corresponding Statement in the Questionnaire
A NPO exists to achieve social outcomes; social profit should be at the helm.	We support and provide community and social services that effectively meet the needs of those we serve.

The 40 statements corresponding to the 40 operational descriptions of the nine PM dimensions (constructs) formed the key items of the quantitative survey questionnaire. Each statement invites a quantitative score from the respondent, based on the Likert type scale being used to rate (indicate the level of agreement to the statement) each statement. The composition of the questionnaire, the Likert type scale being used, validation, testing and the survey administration have been described in section 4.6.2 in Chapter Four. The following subsections describe the content of the questionnaire items that were used to operationalise each PM dimension (construct).

6.6.1. Questionnaire Items that Capture Mission

The Mission is the heart of NPOs since these organisations are *mission driven* organisations. The Mission of a NPO emphasises the reasons for the existence of the organisation categorically. Four questionnaire items capture the Mission as four operational descriptions are associated with Mission (Table 6.4). These four questionnaire items reflect two subdomains: *Social Profit* and *Organisational and People Excellence*. The main aim of the Mission of a NPO is social profit, which

pertains to providing benefits to society. Social profit is created when the lives of targeted group/society is improved (Emerson, Wachowicz, & Chun, 2000). Questions 1 and 5 (i.e. Q1 and Q5) of the questionnaire (Part two) capture social profit, while Q9 and Q13 capture Organisational and People Excellence.

6.6.2. Questionnaire Items that Capture Strategy

Five questionnaire items capture strategy (Table 6.4). These questionnaire items represent two subdomains: *Strategy Planning* (Q16 and Q17) and *Strategy Implementation* (Q18, Q24 and Q39). The questionnaire items under Strategy Planning capture the things the senior leaders do to plan the strategies. The questionnaire items under Strategy Implementation cover the things an organisation does to implement its strategies.

6.6.3. Questionnaire Items that Capture People

Three questionnaire items capture the People dimension. The questionnaire items Q19 and Q29 capture leadership direction while Q30 captures staff training for a high performing work culture.

6.6.4. Questionnaire Items that Capture Organisational Infrastructure

Two questionnaire items capture the Organisational Infrastructure. The questionnaire item Q34 captures the IT and communications infrastructure while Q20 captures marketing and promotion to create *brand awareness*.

6.6.5. Questionnaire Items that Capture Financial Health

As mentioned earlier, the qualitative data indicated that the financial health (wellbeing) of healthcare NPOs depends on two key processes: processes directed at *increasing cash inflows* and processes directed at *monitoring and controlling cash outflows*. In addition, making the *right investments* in IT/infrastructure was found to be necessary to improve financial health (see Table 6.4). The questionnaire items Q3 and Q12 capture relationships with key stakeholders who provide cash inflows: funders, donors, and sponsors. These relationships include providing them information and feedback on how funding was sourced to achieve the mission and objectives of the organisation. The

questionnaire items Q4 and Q14 also capture the cash inflow aspect through fundraising projects and continuous innovation to create new revenue streams. The questionnaire item Q15 captures monitoring and controlling cash outflows through reducing waste and termination of activities that do not result in serving the customer; in a for-profit organisation, such activities could be referred to as non-value adding activities. Finally, Q2 captures investments in the IT/infrastructure.

6.6.6. Questionnaire Items that Capture Processes

As mentioned earlier, the qualitative data indicated that the processes put in place by healthcare NPOs to achieve stakeholder satisfaction can be divided into three subdomains: *Continuous Improvement*; *Design of Safe, Efficient and Effective Processes* (which can be continuous or big-step process improvements); and *Designing the Infrastructure, Technology and Material* to create the necessary support processes to achieve stakeholder satisfaction (see the eight operational descriptions of Processes shown in Table 6.4). The eight questionnaire items that capture Processes represent the three Process sub-dimensions mentioned above as follows. The questionnaire items Q10 and Q40 capture continuous improvement. The questionnaire items Q36, Q23, and Q25 capture safe, efficient and effective processes needed to satisfy the stakeholder requirements. The questionnaire items Q37, Q35 and Q31 capture the extent to which the HR, finance and IT systems have been designed to support the continuous improvement—and, where relevant, big-step improvement—processes.

6.6.7. Questionnaire Items that Capture Client Satisfaction

As mentioned earlier, the qualitative data confirmed that the three stakeholders of NPOs are the Clients, the People, and the Donors. The data also confirmed that each of these stakeholders have their own set of expectations. Further, most of the healthcare NPOs were found to be dealing with clients/patients with various disabilities (e.g. disabilities associated with the very elderly and mental disabilities). The patients (external customers) as well as their family/caregivers invariably expect high quality support and services to the patients (external customers). The questionnaire items Q6, Q8 and Q26 capture the prevalence of the external customer audits and service quality reviews. The questionnaire item Q32 captures the requirement of having to have efficient and effective services that the organisation provides to their external customers.

6.6.8. Questionnaire Items that Capture People Satisfaction

The qualitative data confirmed that even in a nonprofit context satisfaction of the organisation's people (paid employees and volunteers) is an essential requirement for the sustenance of the organisation. Some NPOs were found to be facing difficulty in attracting skilled people. The NPOs were found to be striving for the right balance between monetary and non-monetary rewards for their people by providing opportunities for their development and growth and fair remuneration. In addition the NPOs seem to recognise the performance of their people, allowing them to be involved in planning the high level activities of NPOs. The questionnaire items Q28, Q41, Q21, Q33, Q7 and Q38 capture these people related aspects.

6.6.9. Questionnaire Items that Capture Donor Satisfaction

The qualitative data also found that NPOs endeavour to achieve donor satisfaction (which is necessary for the financial sustainability of the organisation) by giving them the confidence that their moneys are directed towards creating a well performing and an empathic social service to the society. Typically, NPOs seem to be achieving this objective (i.e. donor satisfaction) by: (a) designing their services around the needs of the community (or targeted group); (b) developing and implementing new service models; and (c) collaborating with other organisations to fulfil the responsibility to the community (intended target group). The questionnaire items Q27, Q22 and Q11 capture these donor related aspects.

6.7. THEORETICAL AND PRACTICAL IMPLICATIONS

6.7.1. Theoretical Implications of the Case Study Findings

The purpose of the case study was to gain theoretical and practical knowledge of what set of dimensions constitute an integrated PM system for NPOs and how each dimension could be operationalised. The conceptual model consisting of six PM domains (Figure 4.4) forming tentative theoretical propositions was the basis of the case study. The case study enabled the researcher to collect qualitative data to understand what each PM dimension captures.

The case study corroborated the six PM domains of the conceptual model but the researcher was able to subdivide some PM domains resulting in a PM system consisting nine domains: Mission, Strategy, Financial Health, People, Organisational Infrastructure, Processes, Client Satisfaction, People Satisfaction, and Donor Satisfaction (Figure 6.4). These PM domains were identified as separate themes (or in some cases subthemes) that emerged from the qualitative data—that is responses given by the respondents belonging to the nine case study organisations (Table 6.3). Some PM domains consisted of subdomains and the nine PM domains in effect represented sixteen subdomains. The PM domains and their subdomains were interpreted from a measurement perspective in section 6.6.

One PM dimension that indicated a clear division of divergent subthemes is Stakeholder Satisfaction. As mentioned earlier, Stakeholder Satisfaction represents the needs of three distinct groups having *divergent* needs, which mean that from a measurement perspective, Stakeholder Satisfaction represents three subdomains: Client Satisfaction, People Satisfaction and Donor Satisfaction. Even though LeRoux (2009, p.163) asserts that “resource dependence theory forms the basis for a competing model of stakeholder management in which organisations can and do strategically place some stakeholders interests over others because financial performance (revenue growth) is contingent upon such a strategy” the study revealed that NPOs should pay attention to the requirements of all three stakeholders.

Another PM dimension that indicated a clear division of distinguishable subthemes is Organisational Capability, Infrastructure, and People Development. The qualitative data implied that People as a resource can be distinguished from the Organisational Infrastructure (the latter, for the most part, represents IT). The researcher also observed that People and Organisational Infrastructure sub-dimensions tally with the Human Capital and Information Capital subdomains belonging to the Learning and Growth dimension of the BSC (e.g. see Kaplan & Norton, 2004).

On examination of the theoretical relationships between the nine PM dimensions (Figure 6.4), the reader will observe that different aspects of the stakeholders (more noticeably People and the Donors) appear on either side of Processes. For example the researcher’s model includes a Financial Health dimension as well as a Donor

Satisfaction dimension. The Financial Health dimension represents the processes managers put in place to win donors. These processes are aimed at attracting more funding (from the donors) as well as curb spending. In the nonprofit BSC these processes would be subsumed under Internal Business Processes. The financial perspective of the BSC traditionally represents metrics that satisfy the donor and therefore, this perspective can be treated as being analogous to the Donor Satisfaction PM dimension in the researcher's model. Unlike the researcher's model, the nonprofit BSC does not explicitly model how managerial process lead to Donor Satisfaction. Also, the BSC (the nonprofit version or the regular version) assumes that the *organisational learning* notion embodied in the Learning and Growth perspective results in requisite outcomes for the donors (the Financial perspective) and the clients/customers (the Customer perspective). The BSC does not explicitly capture how a NPO performs in the People dimension. An organisation may take several initiatives to become a learning organisation, which in theory, should reward the People of the organisation intrinsically as argued by Herzberg (1965) in his *two factor theory*. However, in practice, the initiatives to create a learning organisation could result in dissatisfaction of the People, which could have catastrophic effects on the organisation. A direct check on people satisfaction enables managers to take corrective action more readily. The researcher argues that by augmenting the nonprofit BSC, the researcher makes the nonprofit BSC more interpretable both from a theoretical and a practical standpoint.

6.7.2. Practical Implications of the Case Study Findings

The researcher observed that in generating the operational descriptions of PM dimensions and their sub-dimensions (Table 6.4), the researcher was able to identify the critical ingredients or factors that are required to ensure strategic success of NPOs. In performance management, such factors are known as *critical success factors (CSFs)* (Barrett, Balloun, & Weinstein, 2012; Oakland, 2004; Parmenter, 2015). The researcher found that there is one-to-one correspondence between the 40 operational descriptions provided by the researcher (Table 6.4) and the CSFs, because each operational description evolved from a certain critical aspect (a critical factor) of strategic performance, although the CSFs themselves have not been explicitly stated and labelled in the previous sections.

For the convenience of the practitioner, the researcher has labelled these CSFs now, in Table 6.6. Healthcare managers in NPOs can use these critical success factors to identify the suitable key performance indicators (KPIs) for their organisations, given their strategic objectives. The researcher has included a list of KPIs for each CSF, based on the responses given by the respondents of the case study organisations. The list of KPIs is not exhaustive and the KPIs listed by the researcher must be interpreted with caution by the practitioner (the CSF is more robust than the KPI/s). The researcher observes that quite a number of KPIs suggested by the researcher have been mentioned in the literature (e.g. Grigoroudis et al., 2012; Gurd & Gao, 2007; Parmenter, 2015).

Table 6.6 : The Critical Success Factors and Suggested KPIs

PM Dimension	Sub-dimension	Critical Success Factors	Suggested KPIs based on Fieldwork*
Mission	Social Profit	Providing effective community services	Project index
		Improving the living of the community	Number of ongoing programmes Number of improvements made to existing services
	Organisational and People Excellence	Developing people and teams to follow the organisation's philosophy	Number of qualified employees available (FTE)
		Being seen in the community as the leader of the choice	Client satisfaction index (Provider of choice)
Strategy	Strategy Planning	Considering triple bottom line elements in strategy planning	Number of initiatives suggested based on the last survey
		Senior leadership involvement in strategy planning	Management/Professionals experience (person years)
		Collaborating with sector organisations to achieve social responsibility	Number of joint projects with other sector organisations
	Strategy Implementation	Cascading individual and departmental operating plans to each employee	Number of staff workshops on the strategic plans

PM Dimension	Sub-dimension	Critical Success Factors	Suggested KPIs based on Fieldwork*
		Considering organisational structure and resources in strategy implementation	Number of assessments on organisational structure and resources
People	Board of Directors, Professionals and the Senior Leadership	Strengthening the senior leadership team	Number of qualified managers
		Delegating decision making	Number of new strategies for fund raising
	Training	Creating a high performing work culture	Number of training programmes Percentage of trained (qualified) employees
Organisational Infrastructure	Infrastructure and Technology	Developing and implementing paper less initiatives	Percentage of budget Investment in IT and the communications infrastructure
	Marketing & Promotion	Creating brand awareness and public awareness	Number of awareness campaigns (brand promotions)
Financial Health	Monitoring and Controlling Cash Outflows	Cost reduction	Percentage reduction in operating expenses over the previous period
		Investing in infrastructure and IT	ROI on infrastructure and IT
	Monitoring and Increasing Cash Inflows	Providing regular information and feedback to the key stakeholders	Number of publications
		Maintaining good relationships with the key stakeholders	Number of engagements with the donors Donor satisfaction index
	Right Investments	Identifying and introducing new services	Number of new revenue streams
Processes	Continuous Improvement	Continuous improvement of processes for donor outcomes	Number of improvements suggested for existing services based on donor suggestions
		Continuous	Number of improvements

PM Dimension	Sub-dimension	Critical Success Factors	Suggested KPIs based on Fieldwork*	
		improvement of processes for customer outcomes	suggested for existing services, based on client surveys	
		Developing, reviewing and updating the strategies periodically, based on stakeholder feedback and the state of the environment	Number of ongoing reviews on costs and revenues	
		Reviewing the operational strategies based on PM	Number of reviews on operational strategies	
	Infrastructure, Technology & Materials	Developing the organisational infrastructure and technology to deliver quality services to the clients	Number of currently developed IT services	
	Well Designed, Safe, Efficient & Effective Processes	Designing key support processes	Number of cost saving projects	
		Identifying key value-adding processes for strategy implementation	Number of resource allocation plans	
		Designing service delivery processes to satisfy stakeholder requirements	Number of initiatives completed from the most recent survey	
	Client Satisfaction	High Quality of Support & Services	Conducting service audits to review the quality of services	Number of internal and external audits
			Measuring the quality of life of the clients	Number of quality of life issues identified from the most recent audit
			Obtaining clients' feedback	Number of complaints
Measuring the satisfaction of the clients			Client satisfaction index	
People Satisfaction	Skilled & Recognised People	Providing a fair remuneration to paid employees	Employee turnover rate	

PM Dimension	Sub-dimension	Critical Success Factors	Suggested KPIs based on Fieldwork*
		Regular investment in the employees	Number of staff development programmes
		Investing in volunteer development s	Number of training programmes for volunteers
		Communicating strategic plans with the employees	Number of staff workshops on the strategic plans
		Rewarding employees for their performance	Percentage of performance reviews completed on time
		Measuring the satisfaction of the employees	Employee satisfaction index
Donor Satisfaction	High Quality of Support & Services	Designing work/ services around the needs of the community	Percentage of closed cases
		Developing and implementing new models of services	Number of new service models
		Collaborating with other organisations to fulfil social responsibility	Number of collaborations/ joint projects
* Unless stated otherwise, the KPIs are for the review period (e.g. quarter)			

Parmenter (2015) argues that critical success factors are the key to identifying effective KPIs. The researcher agrees with Parmenter, based on her experience in the case study. However, it is also important to note that an organisation may not use a KPI for each CSF listed by the researcher and most certainly, the researcher has not mentioned the KPIs as prescriptions for the practitioner (they are for indicative purposes only). Managers are likely to identify more appropriate KPIs for their own organisations, based on the CSFs listed by the researcher. However, the researcher contends that the researcher's model (operational descriptions and CSFs included) provides more direction to the practitioner than the nonprofit BSC as to how performance measures should be developed for each performance dimension. It is well documented in the literature that the BSC does not provide much guidance to the practitioner as to how CSFs could be identified, which is central for generating KPIs (Anand et al., 2005;

Antonsen, 2014). The onus is passed on to the manager to develop strategy maps, which in practice requires significant external help from consultants. A culture of an organisation may not favour substantial external involvement and/or NPOs in particular may view hiring external consultants as cost prohibitive. The researcher argues that by augmenting the nonprofit BSC (this includes accommodating practitioner concerns), the researcher makes the nonprofit BSC more interpretable and responsive both from a theoretical and a practical standpoint.

6.8. CHAPTER CONCLUSION

In this chapter, the researcher presented the results of the qualitative phase of the study (the case study). The qualitative data collected from the respondents belonging the case study organisations enabled the researcher to elevate her conceptual model into a testable theoretical model (Figure 6.4) consisting of nine PM dimensions. This model is testable on two counts. Firstly, the qualitative data enabled the researcher to operationally describe each PM dimension (Table 6.4). These operational descriptions enabled the researcher to operationally define each PM dimension, by being able to develop a quantitative questionnaire to capture each PM dimension. Secondly, the qualitative data enabled the researcher to posit a definitive set of PM dimensions (nine of them) hypothesised to be causally related, and because each PM dimension (construct) can be captured quantitatively, the hypothesised causal relationships themselves become testable. The model represents—with a greater deal of clarity—the theory underlying the nonprofit BSC. The findings in this chapter answers the first two research questions because the findings answer what set of PM dimensions constitute an integrated PM system for NPOs (RQ1), and what the operational definitions are for each PM dimension (RQ2). Consequently, the researcher was able to achieve the following main outcomes:

1. Understand how the BSC's performance dimensions (domains) should be augmented and operationalised to predict and explain the strategic performance of NPOs; and
2. Develop a PM instrument to test the validity of the BSC's performance dimensions (as augmented).

The first main outcome mentioned equates to achieving the researcher's *first research objective*. The second main outcome mentioned equates to achieving the researcher's *second research objective*, once the PM instrument has been scientifically validated using the quantitative data collected during quantitative phase of the study (see Chapter Eight).

The researcher also discussed the implications of the findings of the case study from a theoretical and practical standpoint (section 6.7). In particular, the researcher explained how adding Financial Health as a separate PM dimension (construct) increases ones theoretical understanding of the nonprofit BSC. Likewise, the researcher also explained why, unlike the BSC, the researcher has included a separate PM dimension labelled 'People Satisfaction'. The researcher also showed how the qualitative data enabled the researcher to discover the CSFs, which help managers to develop KPIs for their own organisation (a set of KPIs have also been suggested for each CSF). The next chapter (Chapter Seven) displays and discusses the descriptive statistics of quantitative data collected through the quantitative questionnaire.

CHAPTER SEVEN

DESCRIPTIVE STATISTICS

The aim of art is to represent not the outward appearance of things, but their inward significance

Aristotle

7.1. INTRODUCTION

The descriptive statistics of the data collected in the quantitative data collection phase (the quantitative survey questionnaire) are presented and discussed in this chapter. The content discussed in this chapter facilitates the discussions provided in the next chapter.

7.2. STATISTICS OF THE QUANTITATIVE STUDY

Out of the 1550 senior managers (930 from Australia and 620 from New Zealand) who were invited to participate in the survey, 232 (15%) responded. This comprised of 121(13%) responses from Australia and 111(18%) responses from New Zealand (Figure 7.1). A higher New Zealand response rate could probably be attributable to the fact that the survey was conducted by a researcher belonging to a New Zealand university.

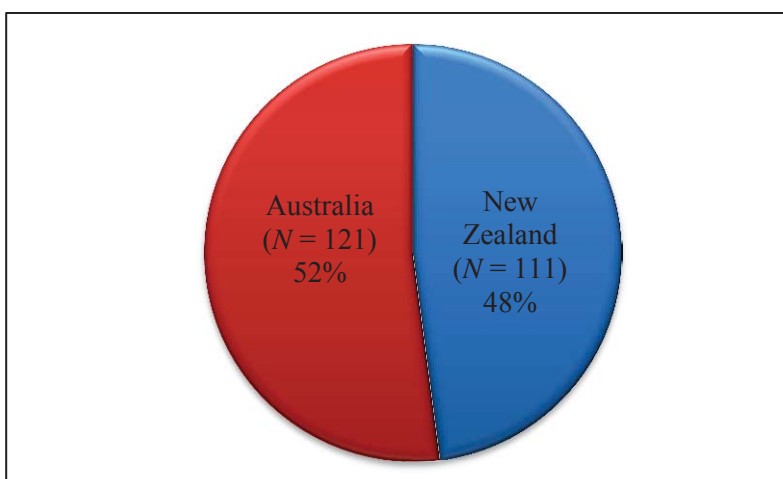


Figure 7.1: Survey representation by country

7.2.1. Response Pattern During the Survey Period

Figure 7.2 depicts the response pattern during the survey period. As evidence in Figure 7.2, the data were collected in two stages.³⁸ The second stage was used purely to increase the number of cases (responses) above 200 for reliability purposes, even though the power analysis indicated that the minimum sample size (data pooled from both countries) required for the study is 96 (section 4.6.1). The first stage of sampling lasted 12 weeks: 8th November 2014 through to 5th February 2015 (a longer duration was allowed due to the summer holiday season in the region). The second stage of sampling lasted two weeks: 26th August 2015 through to 10th September 2015. In the first stage, a high response rate was observed in the first four weeks of the survey (8th November 2014 through to 3rd December 2014). This might be attributable to the fact that the researcher solicited the recipients to respond to the questionnaire within two weeks of receipt of the invitation to participate. A second peak was observed (see Figure 7.2) from 11th January 2015 through to 5th February 2015, which could be attributed to the fact that a reminder was sent out to the original recipients after the summer holidays. The third and final peak shown in Figure 7.2 (26th August 2015 through to 7th September 2015) correspond to the second stage of sampling, which resulted in 25 additional observations.

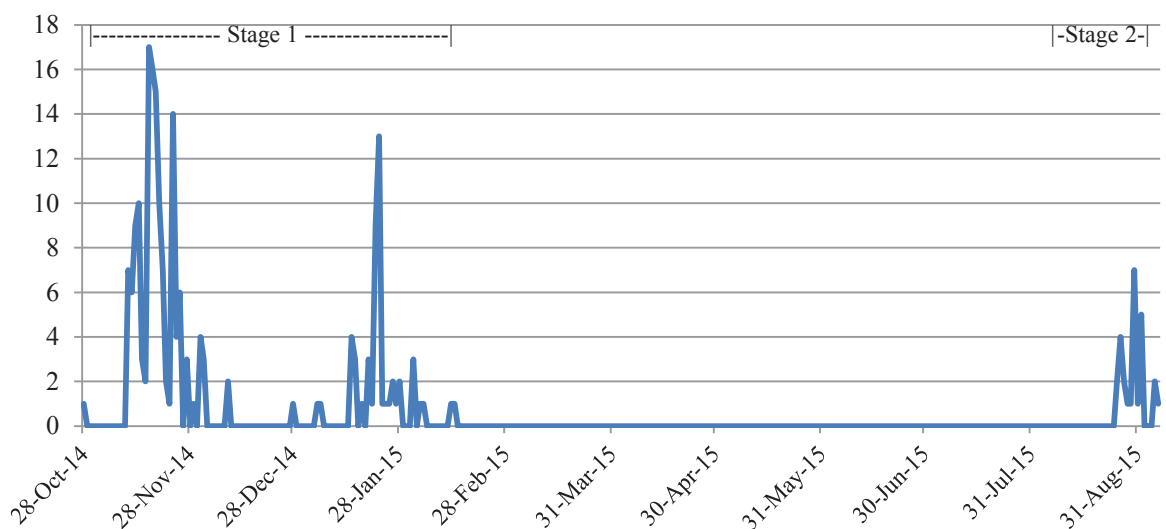


Figure 7.2: Response rate during the survey

³⁸ This does not mean that the New Zealand sample and the Australian sample were “snowball samples”; a snowball sample is a nonprobabilistic sample that is built gradually (like a snowball) when it is difficult to develop probabilistic samples (Cooper & Schindler, 2014).

7.2.2. The Profile of the Survey Participants

Figure 7.3 depicts survey participants (respondents) by designation. Of the 232 who responded, 72 were CEOs (31%), 66 were senior executives (28%), 18 were directors (8%), while the remaining 76 (33%), categorised as “other”, were senior managers responsible for performance monitoring. These results are not surprising because as in the qualitative study, the researcher invited only the senior managers (two from each NPO) to participate in the quantitative survey.

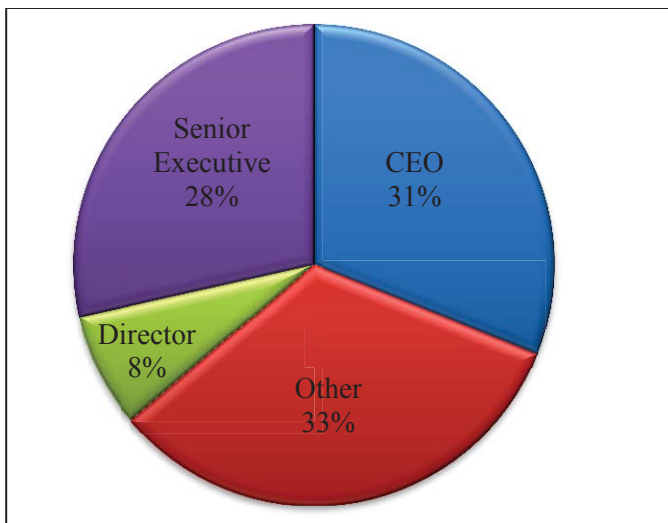


Figure 7.3: Participants by designation

7.2.3. Sector Representation

Figure 7.4 depicts respondents’ organisation by the relevant health subcategory.³⁹ The subcategory classification for coding (the first survey question, as shown in Appendix C) was based on the definition provided by Salamon and Anheier (1997). This definition is also used in the International Classification of NPOs (ICNPO). In keeping with the ICNPO, the category “other” was included as a response category to cover all healthcare services that are not covered in the remaining four subcategories. Figure 7.4 indicates that quite a number of organisations (41%) have declared themselves as organisations providing “other” health services. This does not

³⁹ Since the researcher invited two participants from each organisation to respond and it is not possible to verify how many organisations (if any) were represented by how many respondents (i.e. one or two?), it is not possible to report the actual number of organisations who participated in the survey in Figure 7.4. The same reason applies to Figures 7.5 through to 7.8.

necessarily reflect a nonresponse bias.⁴⁰ From the fieldwork in the qualitative phase of the study (the case study) and in pilot testing of the quantitative questionnaire, the researcher found that some organisations prefer to classify their service as “other”, when in fact their service belonged to one of the main (the first four) categories specified in the ICNPO. Therefore, the proportion under the category “other” in Figure 7.4 is probably a highly inflated figure.

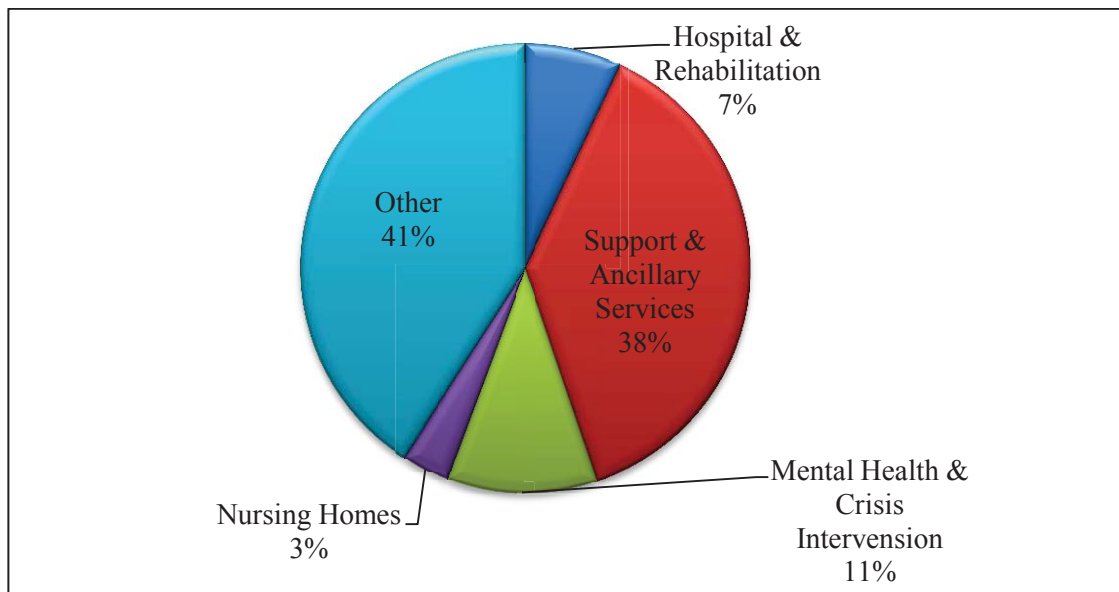


Figure 7.4: Health subcategory wise responses

7.2.4. The Participants’ Familiarity with Performance Measurement Systems

Figure 7.5 depicts the proportion of respondents that use performance measurement (PM) systems. Figure 7.5 indicates that more than half of the respondents (53%) currently use a PM system (an in-house one or an established one such as the BSC) to monitor the performance of their organisation (the corresponding survey question had three categories: “Yes”, “No”, and “Unsure”).

Figure 7.6 provides information on respondents’ familiarity with PM systems. Figure 7.6 indicates that majority of the respondents (80%) have been familiar (“Very Much”

⁴⁰ Unfortunately, the published descriptive statistics on healthcare NPOs in Australia and New Zealand are not classified based on the ICNPO sub-classification. Therefore it is impossible to compare the sub-category proportions of the population and the sample to judge any nonresponse bias. However, it is hard to conceive that as high a proportion as 41% of NPOs in healthcare belong to the “other” sub-category.

or “Somewhat”) with PM systems, whether or not they currently use one. The researcher assumed that familiarity with PM systems is not a prerequisite to qualify as a participant in the survey because the survey does not require specific knowledge (on the part of the respondents) on any specific PM system. The survey was meant to capture certain aspects of strategy monitoring, strategic resource development and usage, good governance, and mission achievement. The senior managers for whom the survey was designed were assumed to be familiar with the area.

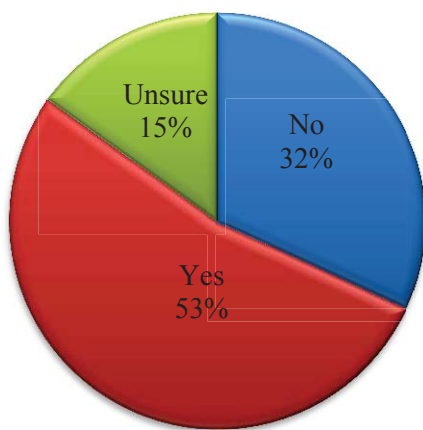


Figure 7.5: Use of a PM system

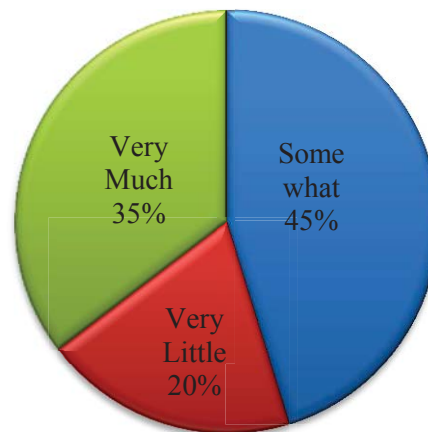


Figure 7.6: Participants' familiarity with PM systems

7.2.5. The Chronological Age of the Responds' Organisation

Figure 7.7 classifies the age of the respondents' organisation. This figure clearly shows that the most of the organisations⁴¹ (73%) have been serving their clients for more than 20 years, suggesting that most respondents in the survey represented mature organisations; this also increases the credibility and usefulness of the survey data.

⁴¹ If a particular organisation has had a predecessor, the respondent was asked to cover the duration of existence of the predecessor also, in counting the chronological age of the present organisation.

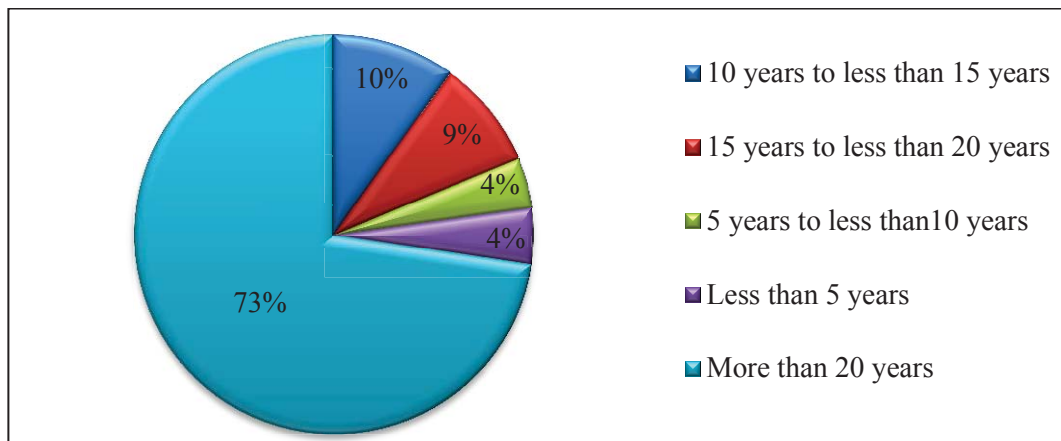


Figure 7.7: Age of the organisation

7.2.6. Funding Streams

Figure 7.8 depicts the level of government funding of the respondents' organisation. Approximately a quarter of the respondents (22%) indicated that their organisation does not receive any government funding, which is bit surprising, given the importance of "health" as a social good. Of the remainder, 30% indicated that their organisations receive some government funding (ranging from 1% to 50% of the revenue), 35% indicated that their organisation receives a sizable amount of government funding (ranging from 51% to 90% of the revenue), while the remaining 13% indicated that they receive more than 90% of their revenue from government funding. These figures indicate that government funding is an important revenue stream for a sizable proportion of NPOs providing healthcare in the community. Participants mentioned the following sources as their sources of nongovernmental funding: philanthropy, community funds, donations, event and fundraising, local councils, industry contracts, sponsors, grants, revenue, and membership fees. It is quite possible that some of the nongovernmental funding streams stated by the managers could have been subjected to *subjective interpretation*. According to Madden and Scaife (2010), the term *philanthropy* is not a widely used term in Australia and New Zealand. Thus it is not clear what exactly the respondents meant by philanthropy. Madden and Scaife assert that the terms "giving" and "charitable giving" are used frequently in Australasia to cover all forms of philanthropy (all type of freely given goods including monetary donations). Thus it is difficult to reliably distinguish between philanthropy and some of the other forms of funding mentioned by the managers. Figure 7.8 suggests that the level of government funding in both countries is similar.

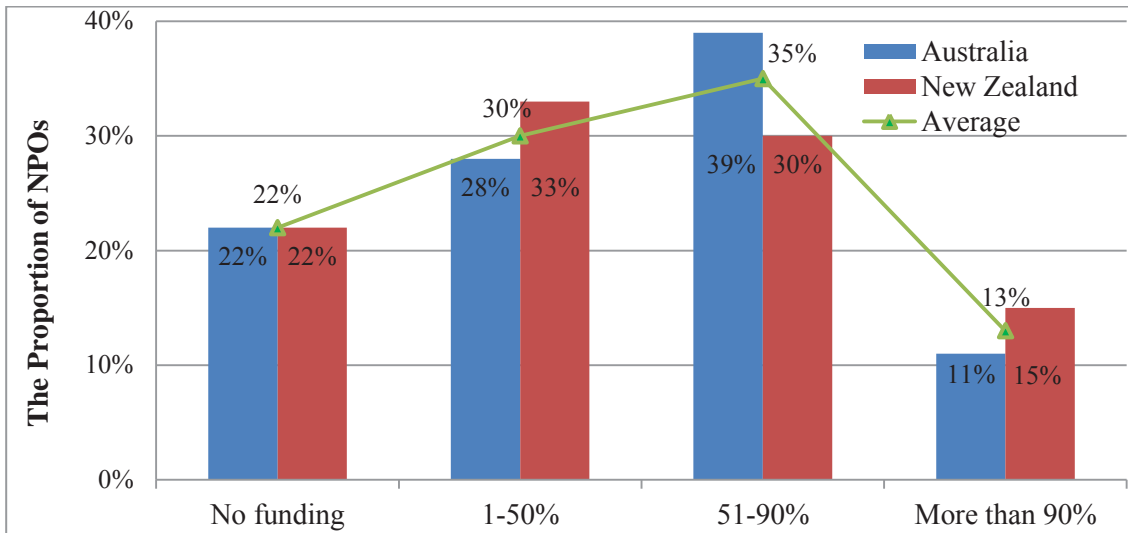


Figure 7.8: Level of government funding

7.2.7. Distribution of Paid Employees

Figure 7.9 depicts the distribution of paid employees within the organisations surveyed. Since the number of paid employees is often used as a proxy to classify the size of an organisation, the information in Figure 7.9 can be used to examine the size of the NPOs that responded to the survey. According to the Australian Bureau of Statistics, the size of an organisation can be classified into one of the following groups, based on the number of paid employees: “small” (20 or less employees), “medium” (more than 20 but less than 200 employees), and “large” (more than 200 employees).

Figure 7.9 shows that 52% of the respondents represented small organisations (i.e. NPOs that employed either no paid employees or 20 or less paid employees). Figure 7.9 also shows that 27% of the respondents represented medium size organisations and the remaining 21% represented large organisations. However, it is important to note that unlike the other two sectors, the nonprofit sector is characterised by a high degree of volunteer labour force participation. Therefore, the number of paid employees may not be a good proxy to categorise NPOs based on their size.

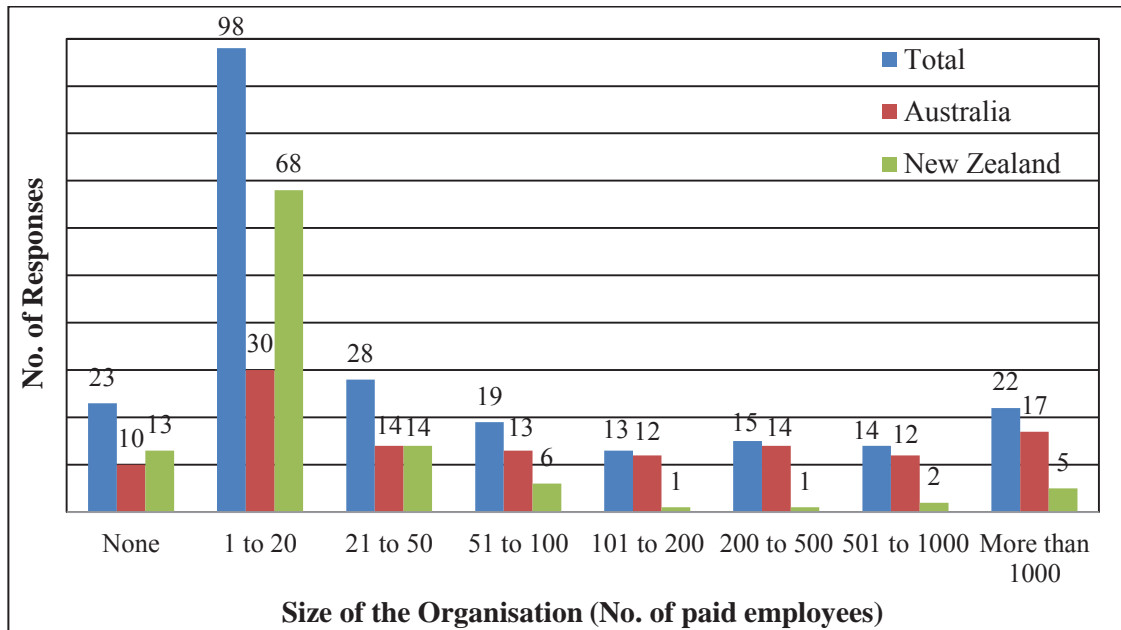


Figure 7.9: Distribution of the paid employees

7.3. ANALYSIS OF INDIVIDUAL SURVEY ITEMS

As mentioned earlier, the statistical analysis of data to test the researcher's hypotheses was conducted using the partial least squares based structural equation modelling (PLSBSEM) approach. This approach is not based on parametric assumptions (the statistical significance of model parameters is examined using bootstrapping which is a nonparametric procedure) and therefore, normality of the observations and even independence, are not required to be tested in PLSBSEM (Chin, 1998). However, the *normal probability plots* of responses for each survey (questionnaire) item indicated that the data are approximately univariate normally distributed; the parametric assumptions become necessary to examine the statistical significance of mean differences and bivariate correlations. Further, there were *no unusual observations* that needed to be removed for statistical analysis. Therefore, there was no necessity to clean the data before the examination of the descriptive statistics of individual survey items and the statistical analysis of data using PLSBSEM. However, since there were some missing data, the researcher used a specific data imputation method (see Chapter Eight) recommend for PLSBSEM, before Likert style responses (part two of the questionnaire) were analysed. Thus, computation of the descriptive statistics of individual survey

items in part two of the questionnaire was conducted via *multiple imputations for missing data*.⁴²

7.3.1. Individual Survey Item Scores

Figure 7.10 shows the mean (average) scores of survey items in the chronological order. These have been calculated from the 223 responses available (9 were excluded as mentioned in section 4.7.1). Five survey items have been colour coded in red to alert the reader that these items were not supported by the data in the statistical analysis used for testing the research hypotheses (details in Chapter Eight). The scope (in brief) of each of the five items that were not supported by data is as follows:

- Q10 Continual process improvement for better outcomes for the clients
- Q 30 Staff training to maintain a high performance work culture
- Q 31 Infrastructure/technology development to improve service quality
- Q 38 Staff workshops on strategic and business plans
- Q 39 Operating plans being traced back to each employee

The mean score of all the survey items except Q3, Q5 and Q20 are below 6.0, in the 1 to 7 Likert scale. To recollect the seven point scale used (details in Chapter Four), 7 indicated a *strong agreement* with the statement meant to cover a particular survey item; 6 indicated an *agreement*; 5 indicated a *somewhat agreement*; 4 indicated an *indifference* (neutral) and so on. The reader should note that the researcher conducted a statistical test to exclude method bias in the responses (details in Chapter Eight).

⁴² The reader should also note that although the researcher has reported measures of central tendency and dispersion in this section to two decimal places, because each survey item is based on a Likert scale (the seven point scale was assumed to be interval), the reported central tendency and dispersion values of the measures should not be treated as if they have come from a ratio scale.

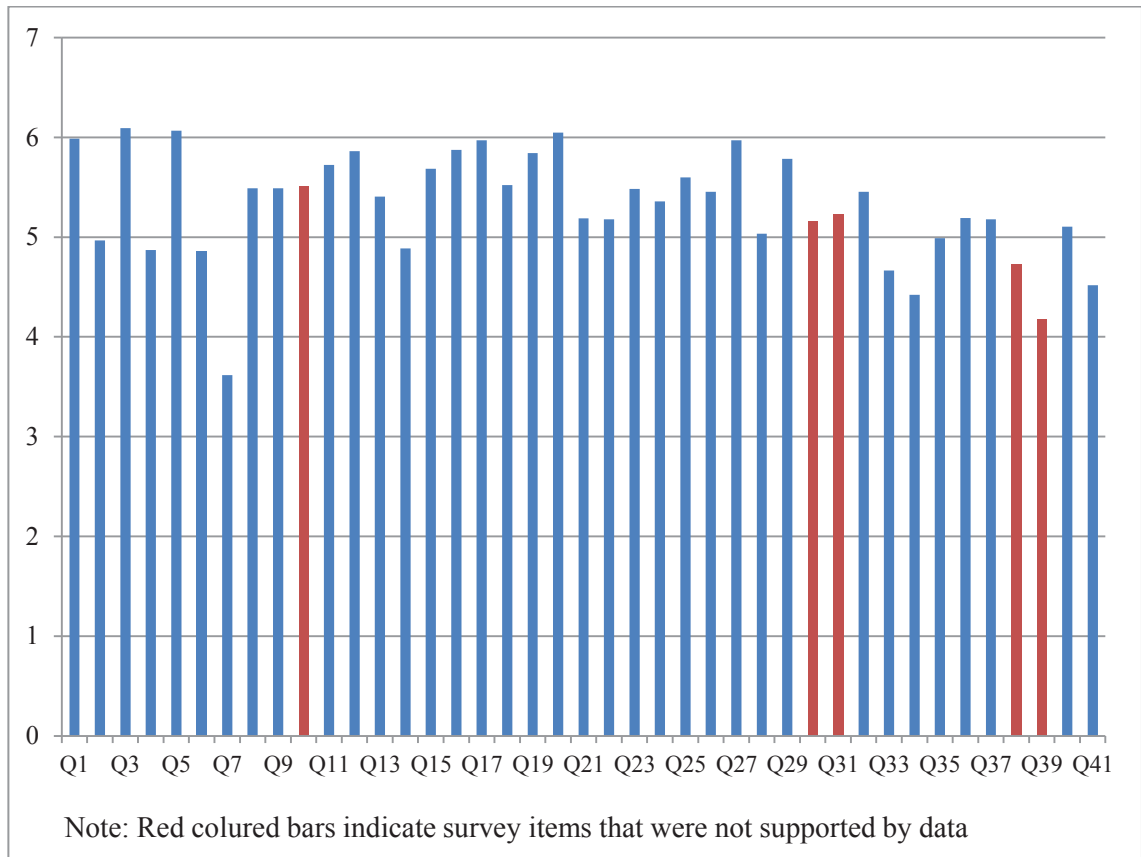


Figure 7.10: Mean scores of survey items in the chronological order

Table 7.1 depicts the measures of central tendency (mean and median) and the dispersion of each item. In Table 7.1, the survey items have been colour coded as follows: red indicates highly scored items while blue indicates weakly scored items. The highly scored items (≥ 6) can be considered as areas of strength for the 223 organisations surveyed (on average) while the lowest scoring items can be considered as areas that need most improvement (on average). Needless to say, each organisation will have its own strengths and weaknesses.

Table 7.1: Measures of Central Tendency and the Dispersion

Item Number	Mean	Median	Standard Deviation	Max	Min
Q3	6.09	6	1.07	7	1
Q5	6.07	7	1.35	7	1
Q20	6.05	6	1.34	7	1
Q1	5.99	6	1.22	7	1
Q17	5.97	6	1.39	7	1

Item Number	Mean	Median	Standard Deviation	Max	Min
Q27	5.97	6	1.06	7	1
Q16	5.88	6	1.23	7	1
Q12	5.86	6	1.25	7	1
Q19	5.84	6	1.31	7	1
Q29	5.78	6	1.34	7	1
Q11	5.72	6	1.14	7	2
Q15	5.68	6	1.38	7	1
Q25	5.60	6	1.39	7	1
Q18	5.52	6	1.31	7	1
Q10	5.51	6	1.41	7	1
Q8	5.49	6	1.53	7	1
Q9	5.49	6	1.31	7	2
Q23	5.48	6	1.42	7	1
Q26	5.45	6	1.58	7	1
Q32	5.45	6	1.46	7	1
Q13	5.41	6	1.43	7	1
Q24	5.36	6	1.28	7	1
Q31	5.22	5	1.39	7	1
Q36	5.19	6	1.52	7	1
Q21	5.19	6	1.59	7	1
Q22	5.18	5	1.43	7	1
Q37	5.18	6	1.60	7	1
Q30	5.15	5	1.56	7	1
Q40	5.11	5	1.50	7	1
Q28	5.03	5	1.59	7	1
Q35	4.99	5	1.45	7	1
Q2	4.97	5	1.65	7	1
Q14	4.89	5	1.59	7	1
Q4	4.87	6	2.09	7	1
Q6	4.86	5	1.77	7	1
Q38	4.73	5	1.74	7	1
Q33	4.67	5	1.78	7	1
Q41	4.52	5	1.79	7	1
Q34	4.42	5	1.78	7	1
Q39	4.18	4	1.84	7	1
Q7	3.62	4	1.88	7	1

Note: As mentioned earlier, Q10, Q30, Q31, Q38, and Q 39 were not supported by the data in the statistical analysis used for testing the research hypotheses.

Table 7.2 depicts the mean scores and the standard deviations of the survey items of the Australian participants and New Zealand participants separately. It is evident that the mean score of questionnaire item 7 is very low for both countries: 3.91 for Australia and 3.28 for New Zealand. In terms of the Likert scale being used, these mean values

approximately translate to “neither agree nor disagree” and “somewhat disagree” for Australia and New Zealand respectively. The questionnaire item seeks the agreement (or otherwise) to the following statement: “*We use recognition schemes to reward our employees for their performance*”.

The low mean score for the above statement implies that on average, the Australian and New Zealand NPOs *do not* agree that recognition schemes need to be used to reward their employees for their performance. This is quite unusual because reward and recognition is regarded as an important requirement to motivate and retain people (Armstrong, 2014; McDermott & O'Dell, 2001; Milne, 2007). Therefore it is necessary to examine the response of a complementary questionnaire item to infer what might have caused respondents not to agree with the aforementioned statement (on average).

Questionnaire item 33, which seeks agreement to the statement “we measure the satisfaction of our employees on a regular basis”, can be considered as a questionnaire item that complements questionnaire item 7. As evidenced in Table 7.2, questionnaire item 33 secured mean scores of 4.82 and 4.48 for Australia and New Zealand respectively; in terms of the Likert scale being used, these mean scores approximately translate to a low level of agreement to “regular employee satisfaction measurement”. One plausible explanation for low level of agreement to the statements given in both questionnaire items (item 7 and item 33) mentioned above is that because the nonprofit healthcare sector is characterised by high levels of voluntary labour force participation rates, and because volunteers do not work for tangible rewards, managers may not see any need in rewarding their volunteers extrinsically. It can be argued that volunteers are intrinsically motivated (seeking psychological rewards) rather than extrinsically motivated (seeking tangible rewards such as money, gifts and other benefits of tangible nature) to perform (Bang & Ross, 2009; Bussell & Forbes, 2002). The researcher wishes to mention that the objectives of the research, and therefore the research design, do not allow her to conduct a substantive analysis on organisational behaviour in NPOs.

Table 7.2 shows that the mean scores for Australia are higher than the corresponding New Zealand mean sources, except in the case of six questionnaire items (highlighted). Similarly, there are greater numbers of highly scored questionnaire items (> 6) in the Australian sample than in the New Zealand sample. However, the reader will note that

the t -values for mean differences are not significantly higher than the critical t -value of 1.97, indicating that the differences between means scores for Australia and New Zealand are unlikely to be of practical significance (possible exception could be questionnaire item 7, which was discussed earlier) to seek any reason for the discrepancies.⁴³ However, in the next chapter, as an extension to her model testing, the researcher has fitted Australian and New Zealand data separately to her model to examine whether or not there is any significant country effect on her hypotheses.

Table 7.2: The Mean Scores Based on Country

Australia			NZ			t -statistic for Mean Difference
Item Number	Mean	Standard Deviation	Item Number	Mean	Standard Deviation	
Q1	6.00	1.27	Q1	5.97	1.18	0.18
Q2	5.08	1.61	Q2	4.84	1.70	1.07
Q3	6.13	0.97	Q3	6.05	1.18	0.49
Q4	5.10	1.97	Q4	4.61	2.21	1.68
Q5	6.04	1.41	Q5	6.09	1.28	-0.26
Q6	4.94	1.72	Q6	4.77	1.85	0.66
Q7	3.91	1.85	Q7	3.28	1.85	2.46
Q8	5.54	1.49	Q8	5.43	1.59	0.48
Q9	5.42	1.37	Q9	5.57	1.24	-0.82
Q10	5.54	1.31	Q10	5.47	1.53	0.31
Q11	5.71	1.05	Q11	5.73	1.25	-0.11
Q12	5.90	1.18	Q12	5.81	1.33	0.5
Q13	5.49	1.30	Q13	5.31	1.56	0.91
Q14	5.03	1.40	Q14	4.72	1.78	1.36
Q15	5.67	1.37	Q15	5.70	1.39	-0.16
Q16	5.93	1.24	Q16	5.81	1.22	0.67

⁴³ The t -value of mean difference exceeds the critical t value for six questionnaire items: Q7, Q19, Q20, Q21, Q26, and Q38. This means that at 5% significance level there is a difference in the means scores of these six measures (Australian mean being greater than the New Zealand mean). To the lay world outside Australia and New Zealand, there is very little difference between the Australian culture and New Zealand culture. However, trans-Tasman studies on culture suggest that Australians are more positive about their performance relative to New Zealanders (Green & Power, 2006), which could translate to a higher mean score for Australia. In any case, a substantive analysis on the effect of any trans-Tasman culture difference on performance scoring is beyond the scope of the study.

Australia			NZ			<i>t</i> -statistic for Mean Difference
Item Number	Mean	Standard Deviation	Item Number	Mean	Standard Deviation	
Q17	6.08	1.34	Q17	5.85	1.44	1.21
Q18	5.66	1.17	Q18	5.36	1.44	1.64
Q19	6.02	1.13	Q19	5.64	1.47	2.07
Q20	6.22	1.22	Q20	5.85	1.44	2.03
Q21	5.41	1.42	Q21	4.93	1.74	2.18
Q22	5.24	1.36	Q22	5.10	1.51	0.69
Q23	5.53	1.32	Q23	5.43	1.54	0.47
Q24	5.41	1.22	Q24	5.30	1.35	0.62
Q25	5.75	1.27	Q25	5.42	1.51	1.68
Q26	5.66	1.54	Q26	5.22	1.60	2.04
Q27	5.96	1.07	Q27	5.99	1.05	-0.23
Q28	5.13	1.53	Q28	4.92	1.64	0.98
Q29	5.92	1.11	Q29	5.63	1.56	1.53
Q30	5.28	1.42	Q30	5.01	1.69	1.22
Q31	5.30	1.35	Q31	5.13	1.44	0.87
Q32	5.42	1.45	Q32	5.49	1.47	-0.037
Q33	4.82	1.72	Q33	4.48	1.83	1.36
Q34	4.55	1.72	Q34	4.27	1.83	1.15
Q35	5.11	1.37	Q35	4.86	1.53	1.24
Q36	5.32	1.39	Q36	5.04	1.65	1.32
Q37	5.38	1.50	Q37	4.95	1.69	1.91
Q38	4.97	1.56	Q38	4.44	1.89	2.19
Q39	4.21	1.78	Q39	4.14	1.92	0.24
Q40	5.20	1.40	Q40	5.00	1.61	0.93
Q41	4.62	1.77	Q41	4.40	1.82	0.86

Figure 7.11 depicts the mean scores of the items based on the participants' familiarity with the PM systems. As evidenced from the figure, in general, the mean scores are higher for the "Very Much" familiar category relative to "Somewhat" familiar category, which in turn is higher than the "Very Little" familiarity category. This implies that on average, the more the organisations that are familiar with PM systems, the more they

tend to rate their performance highly in most of the assessment areas on strategic performance.

Figure 7.12, Figure 7.13, and Figure 7.14 depict the mean scores of the items based participant’s familiarity (“Very Much”, “Somewhat”, and “Very Little” respectively) with PM systems by country. The Figure 7.12 indicates that with the exception of four questionnaire items, the mean scores for Australia are *higher* for the “Very Much” category relative to New Zealand (see the * marks in Figure 7.12). Conversely, with the exception of five questionnaire items, the mean scores for New Zealand data are higher for the “Somewhat” than for Australia (see Figure 7.13). It is easy to show (through the two sample *t*- test) that none of the differences mentioned above are of any practical significance. This implies that on average, participants who are familiar with the PM systems, no matter to which Australasian country their NPO belongs, tend to rate their performance highly in most of the assessment areas on strategic performance.

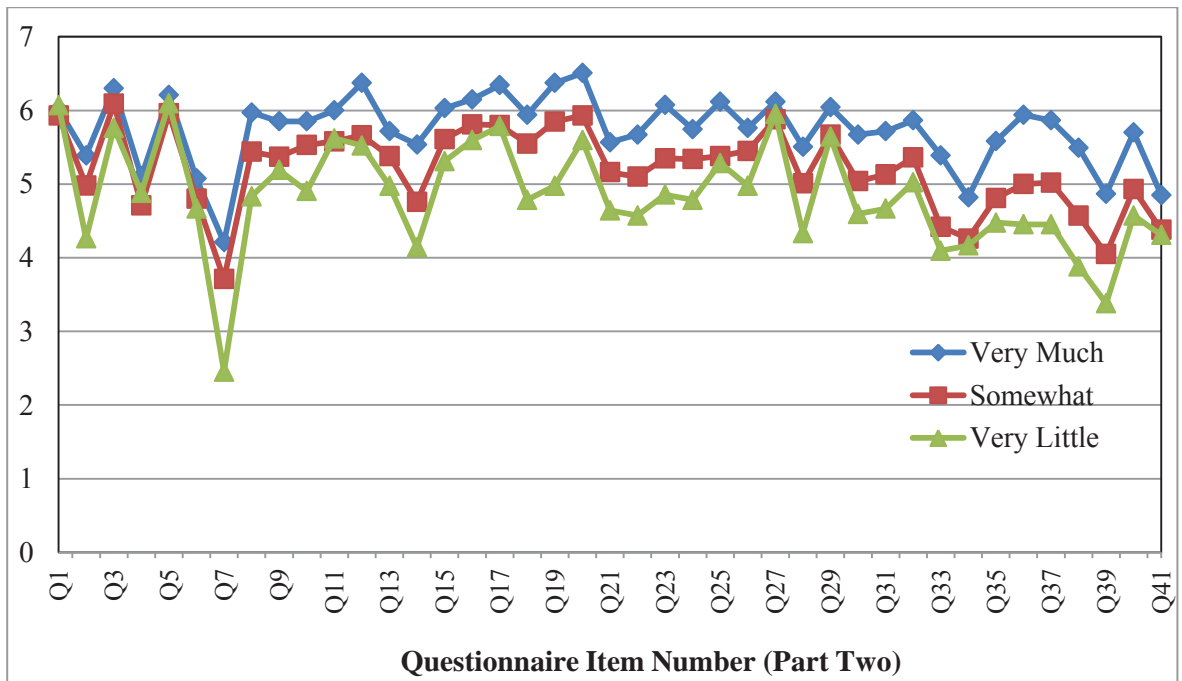


Figure 7.11: Mean scores of survey items based on the participants’ familiarity of PM systems

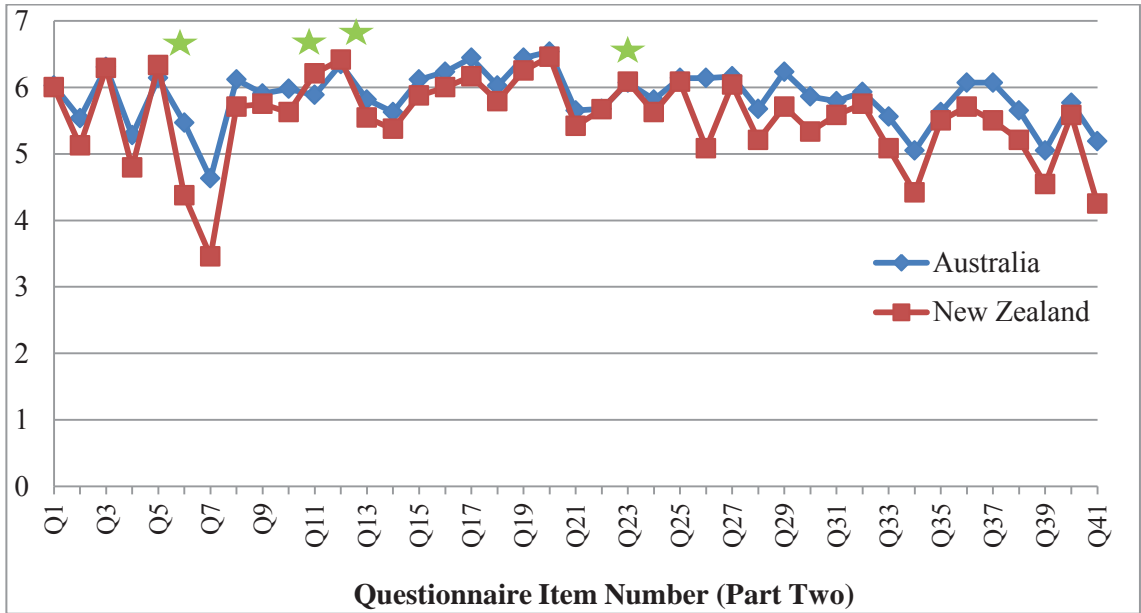


Figure 7.12: Mean scores of survey items by country for participants who rated themselves as being very much familiar with PM systems

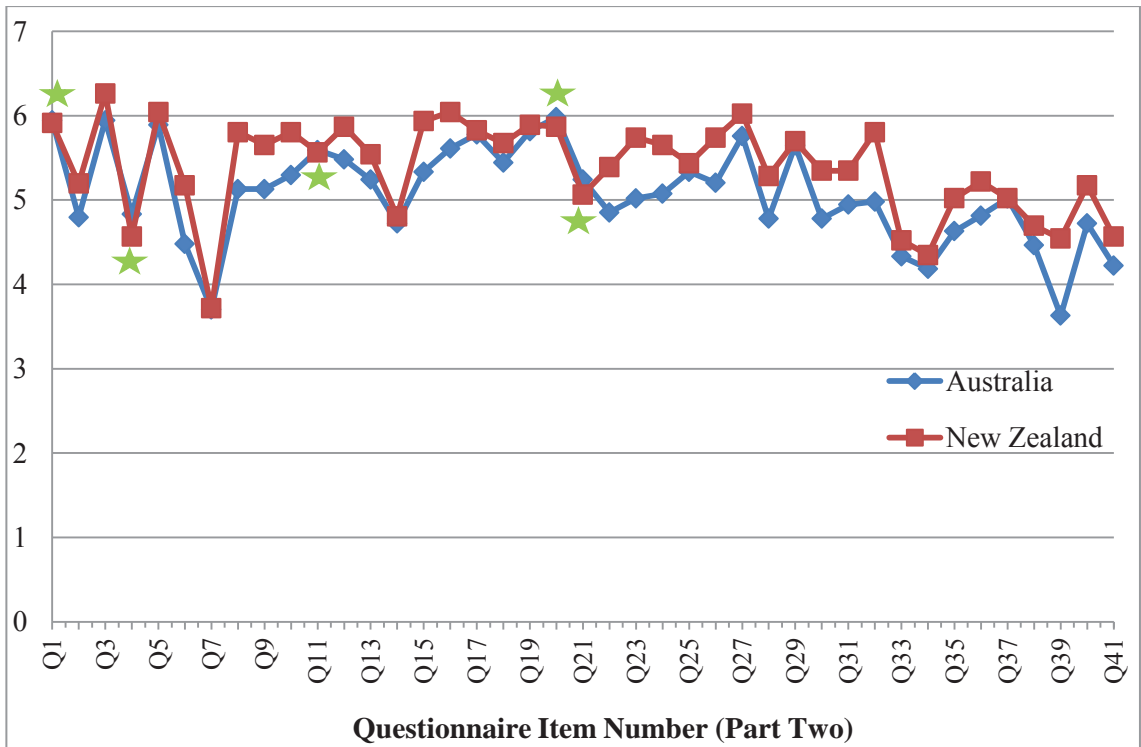


Figure 7.13: Mean scores of survey items by country for participants who rates themselves as being somewhat familiar with PM systems

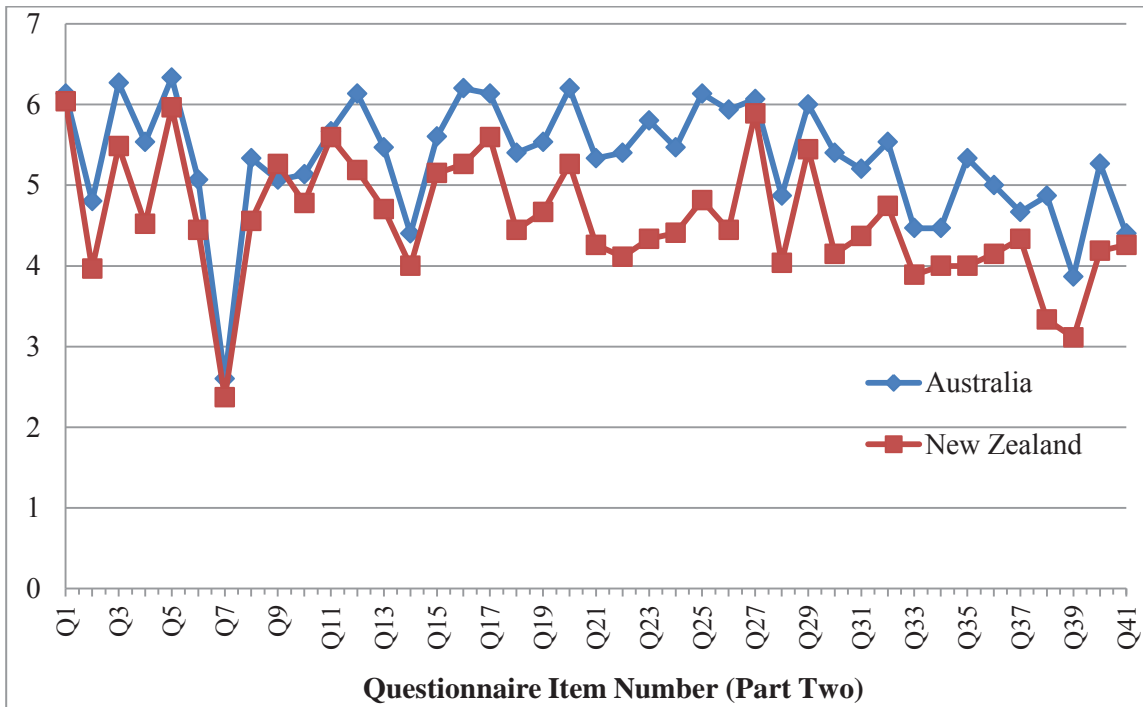


Figure 7.14: Mean scores of survey items by country for participants who rated themselves as being very little familiar with PM systems

7.3.2. Examination of Bivariate Correlations between Questionnaire Items

Table 7.3 depicts the bivariate correlations between questionnaire items. The survey items in Table 7.3 have been rearranged (hence not in the chronological order) so that items belonging to the same construct (performance dimension) becomes adjacent to one another. Since items (measures) belonging to a conventional construct are being viewed as manifestations of the construct, and items belonging to one construct are required to be reliably distinguished from the items belonging to another construct (see Chapter Five for technical details), in general, correlations between items belonging to a particular construct are expected to be stronger (convergent validity) relative to the correlations between items belonging to different constructs (discriminant validity) (Campbell & Fiske, 1959; Chin, 1998; Gefen & Straub, 2005). This requirement is not applicable when the meaning of the construct is formed as a composite of measures—that is, in the case of a formative construct (Diamantopoulos et al., 2008; Petter, Straub, & Rai, 2007). The construct Financial Health is modelled as a formative construct and therefore the conventional psychometric properties such as convergent validity and discriminant validity will not be applicable to Financial Health.

Examination of the correlations of the questionnaire items that belong to the same construct (Table 7.3) indicates that they are strong (with the exception of the items belonging to the formative construct). However, the correlations in Table 7.3 indicates that although the correlations between items belonging to the same construct are stronger than the correlations between items belonging to different constructs (cross-correlations) in general, there are quite a few strong cross-correlations. Prior studies on PM systems also show similar results (e.g. He, Hill, Wang, & Yue, 2011; Jayamaha, Grigg, & Mann, 2011; Saunders & Mann, 2005). Jayamaha et al. argue that relatedness of measures in a PM system is a desired feature when an organisation implements intervention programmes because such programmes are desired to be simultaneously be affecting multiple measures belonging to different performance dimensions—thus implying significant levels of inherent inter-correlations. Chapter Eight provides an extensive analysis of the validity of the measurement scales.

7.3.3. Aggregation of Questionnaire Item Scores to Examine Data at Performance Dimension/Construct Level

The black square bullets in Figure 7.15 depict the average scores at PM dimension (construct) level; the vertical bars in the figure depict the average scores for each questionnaire item belonging to each PM dimension. For the purpose of this chapter, the score of a PM dimension is taken as the arithmetic mean of the corresponding questionnaire item scores, which means that all questionnaire items belonging to a construct have been given equal weight.⁴⁴ The questionnaire items have been colour coded to highlight the PM dimension to which each questionnaire item belongs. As evident from Figure 7.15, the average score for the People Satisfaction dimension is less than the average scores for other PM dimensions. The *F* test indicated that the differences between the means are significant ($F = 22.56, p < 0.001$). The main technical reason for the low average score for the People Satisfaction dimension is that the lowest scoring questionnaire item (Q7- the use of recognition schemes to reward the employees) belongs to this construct.

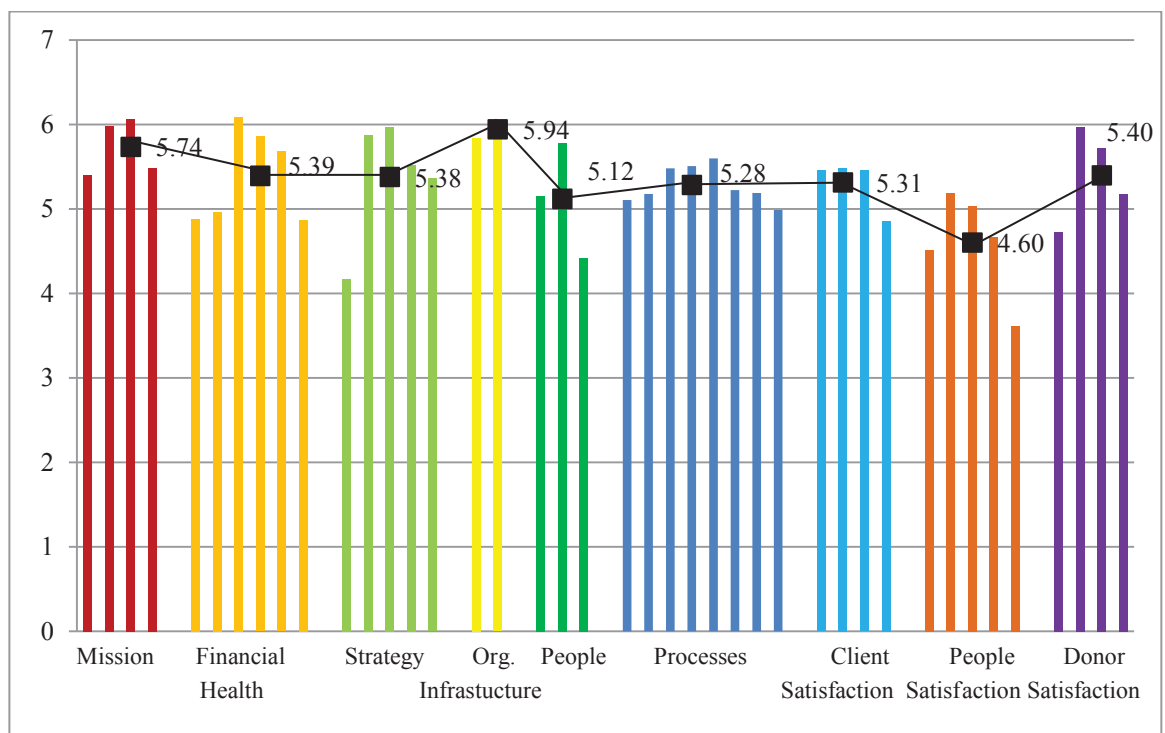


Figure 7.15: Mean scores of survey items based aggregated at performance dimension level

⁴⁴ In PLSBSEM, weights for each item is estimated as part of the parameter estimation procedure to compute the score of the construct as a weighted linear combination of its items.

The square bullet points in Figure 7.16 depict the average scores at PM dimension level for Australia and New Zealand separately; the vertical bars in the figure depict the average scores for each questionnaire item belonging to each PM dimension, by country. As evidenced from the figure, average scores of PM dimensions are higher for Australia than for New Zealand, except for Mission and Donor Satisfaction, for which, both countries have secured the same scores (at 2-dp). However, none of the apparent differences mentioned above were found to be significant, based on the *t*-tests.

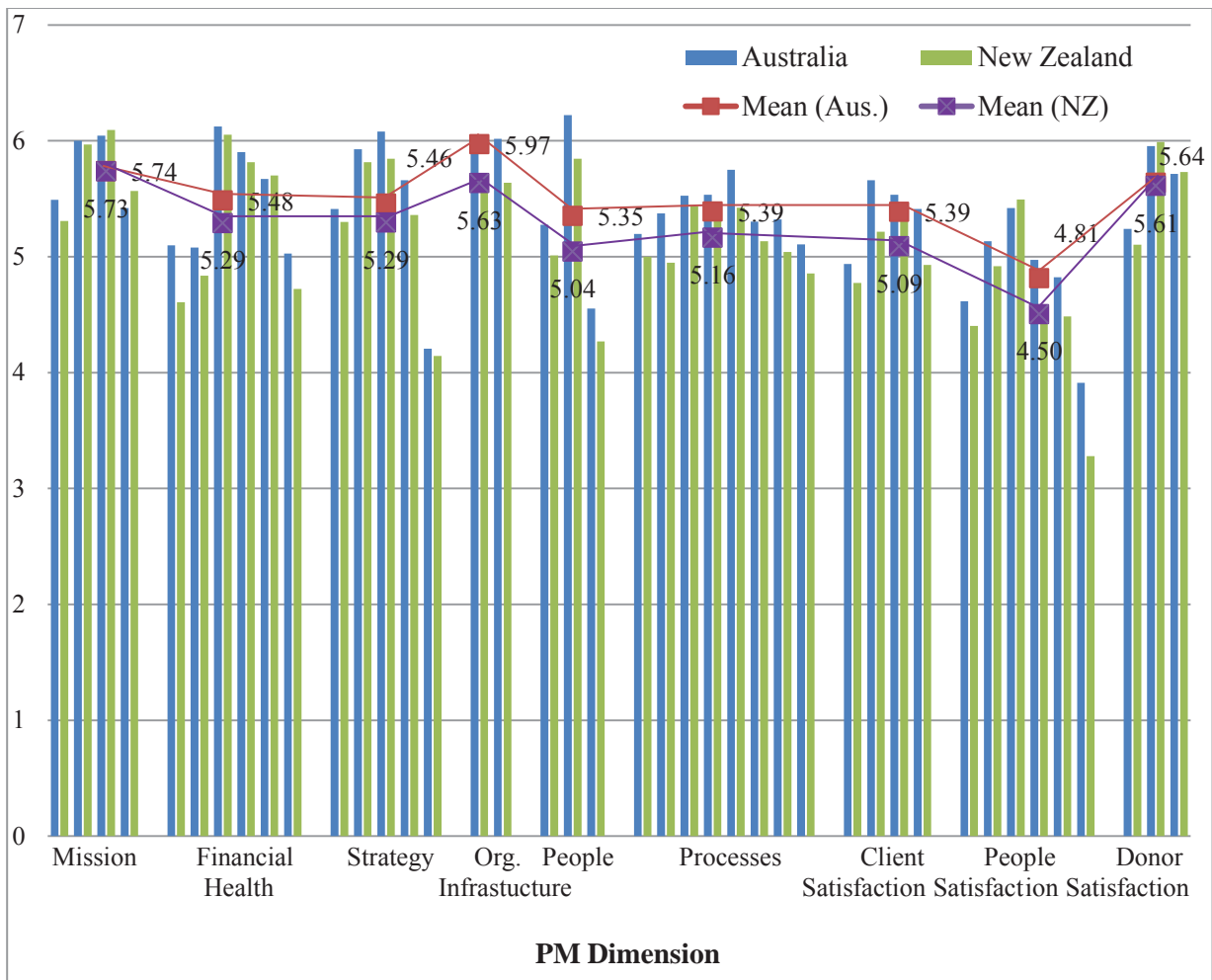


Figure 7.16: Mean scores of survey items aggregated at PM dimension level and by country

7.4. CHAPTER CONCLUSION

The descriptive statistics pertaining to Australian and New Zealand respondents and their organisations were presented in this chapter. An important finding was the similarity of the mean scores between Australia and New Zealand for most of the questionnaire items and certainly for all of the PM dimensions. The next chapter covers the quantitative data analysis and the resulting discussion.

CHAPTER EIGHT

EMPIRICAL TEST RESULTS AND DISCUSSION OF THE THEORETICAL MODELS

Argument is conclusive, but it does not remove doubt, so that the mind may rest in the sure knowledge of the truth, unless it finds it by the method of experiment

Roger Bacon (c. 1214–1294)

8.1. INTRODUCTION

In Chapter Six, the researcher presented the results of the qualitative study (the case study) and discussed the findings. This enabled the researcher to determine what set of PM dimensions constitute an integrated PM system for NPOs (this relates to RQ1); what their operational definitions are (this relates to RQ2); and what theoretical relationships exist between the PM dimensions (this relates to RQ3). Having shown the descriptive statistics of quantitative data in the previous chapter, the researcher presents and discusses the results of the quantitative study in this chapter. The quantitative study was designed to empirically test the researcher's *main theoretical model* (see Chapter Four). In so doing, the researcher validates the PM dimensions (via construct validity), their operational definitions (again, via construct validity), and the hypothesised theoretical relationships between the PM dimensions (via statistical conclusion validity). Therefore, the results and the accompanying discussion provided in this chapter support the findings in the Chapter Six, thus providing more robust answers to RQ1, RQ2 and RQ3. In addition, in testing the *ancillary model* (the measurement model for overall strategic performance) empirically, in this chapter, the researcher estimates the weights that managers should assign to each PM dimension and its subparts, in assessing the overall strategic performance of a NPO. Therefore, this chapter also provides findings to answer the fourth research question (RQ4).

This chapter is organised as follows. Section 8.2 provides the results on measurement validity of the constructs (PM dimensions) of the main model, under four subsections. Section 8.2.1 covers the data screening procedures that the researcher adopted before

proceeding with her data analysis. Section 8.2.2 provides evidence of absence of ‘method bias’ (Podsakoff et al., 2003), a key source of *systematic error* often associated with self-administered surveys. Section 8.2.3 provides evidence of reliability of the measurement scales used for each construct (PM dimension). Section 0 provides evidence of validity of the constructs of the main model (i.e. validity of the operational definitions of the PM dimensions). Section 8.3 provides evidence of statistical conclusion validity associated with the hypothesised relationships between the constructs (PM dimensions) of the main model. As mentioned in Chapter Four (section 4.6.3) and Chapter Five (section 5.4.2), the partial least squares based structural equation modelling (PLSBSEM) technique was used to conduct tests on different facets of validity.

Section 8.4 discusses the results for the main model from a theoretical perspective, while section 8.5 discusses the results from a practical perspective. Thus sections 8.4 and 8.5 address the *third research objective*. Section 8.6 expands the results for the main model by examining the hypothesised relationships by country (i.e. Australia and New Zealand separately). This is to examine whether there seems to be a significant country effect (Australia versus New Zealand) for the theoretical relationships. Section 8.7 covers the results and discussion for the ancillary model, and thereby addresses the *fourth research objective*, by answering RQ4. Section 8.8 offers implications for management practice and performance measurement, which also includes guidelines for nonprofit practitioners to assess strategic performance of their organisations. Thus section 8.8 also achieves the *general research objective* of the study. Section 8.9 provides a comparison between the researcher’s PM models and the BSC (as mentioned earlier, the nonprofit BSC provided the basis for model development) and organisational excellence models. This achieves the *fifth and final research objective*. Finally, section 8.10 concludes the chapter through a brief summary, highlighting the key findings.

8.2. RESULTS FOR MEASUREMENT VALIDITY

8.2.1. Data Screening

As the initial step of data analysis, as mentioned earlier, the data were screened to check whether there have been any data entry errors: unusual characters being entered and/or out of range values having been entered. The only characters that were acceptable were integers from 1 to 7 for data cells containing responses for survey questions relevant to the 7-point Likert scale. Next, the cells were examined to ascertain whether the missing values have been appropriately represented (this facilitates identification of missing data cells and imputation of data for missing data cells), and how many cases (data records/rows) need to be discarded due to excessive levels of missing values.

The problem of missing data is typical in survey research on social and behavioural sciences (Hair et al., 2014). Usually, a case (a data record) is deleted from the data file if the number of missing values corresponding to the case (data record/row) exceeds 15%. Given that the quantitative survey questionnaire contained 41 survey items in the Likert scale, 15% of missing values corresponds to six unanswered questions (survey items) by a particular respondent (case). Thus if any case (data record/row) contained seven or more unanswered questions, these were deleted. In addition, if there were to be responses from a respondent (case) who has provided the exact same level of agreement to all survey items (e.g. a 6, meaning “agree” in the Likert scale) such cases were to be treated as ‘suspect’, warranting ‘case-wise data elimination’. Perhaps unsurprisingly, there was not a single suspect case (a flat score or something close being very flat) warranting deletion based on this criterion. As mentioned earlier (section 4.7.1), 9 cases were deleted, resulting in a usable sample size of 223. The data cells that contained missing values were filled using the data imputation option “median of nearby points” in SPSS (Gerber & Finn, 2013, pp.23-24).

After this screening, data were further examined using IBM SPSS 22 software for univariate outliers (any z value > 3.0) and univariate normality (frequency distribution histograms and normal probability plots in the form of Q-Q plots). There weren't any univariate outliers and the data were found to be approximately univariate normal based on the frequency distribution histograms and the Q-Q plots.

8.2.2. Testing for Absence of Common Method Bias

The common method bias, which is also known as the common method variance, is a systematic error that results from “variance that is attributable to the measurement method rather than to the constructs the measures represent” (Podsakoff et al., 2003, p.879). To ensure that the responses do not suffer from common method bias, as mentioned earlier, the researcher conducted Harman’s single factor test; this test analyses the dimensions of the response items (in a survey questionnaire, all the questionnaire items) to ensure that all the response items taken together, do not resolve into a single dimension (Conway & Lance, 2010; Podsakoff et al., 2003).

According to the Harman’s single factor test, if the principal components analyses (PCA) of the survey items extract only a single factor (component), the responses of a survey are deemed to suffer from common method bias. The PCA analysis of the survey data containing Likert style responses showed that 7 factors (based on the Kaiser’s criterion of Eigenvalues > 1.0) are required to extract the variability of the survey items. The Eigenvalues of these factors (components), for the un-rotated solution were found to be 18.432, 2.149, 1.839, 1.468, 1.324, 1.243, and 1.074. This demonstrated that it is unlikely that the survey has been affected by common method bias.

8.2.3. Scale Reliability

Reliability (more precisely, internal consistency reliability) of a measurement scale is a necessary condition for validity of a conventional construct (Nunnally & Bernstein, 1994), which in the researcher’s main model, applies to all but the formative construct Financial Health. Scale reliability reflects the extent to which measurement items (indicators) of a conventional construct relate to one another (i.e. inter-correlate) in operationalising the construct (Hair et al., 2014; Nunnally & Bernstein, 1994). As mentioned in Chapter Five, the measures that constitute the measurement scale of a conventional construct are mere manifestations of the underlying concept embodied in the construct (Nunnally & Bernstein, 1994). As such, the measures are required to be co-varying (inter-correlate).

There are two indicators that are commonly used to test the reliability: Cronbach's coefficient alpha (α) (Cronbach, 1951) and the composite reliability coefficient (ρ) (Werts et al., 1974). Traditionally, Cronbach's α provides conservative estimates of reliability as the calculation assumes that all measures of a construct are equally reliable, an assumption known as Tau equivalency; for this reason, PLSBSEM researchers (Tau equivalency is not assumed in PLSBSEM) are encouraged to use ρ as the primary measure of reliability (Chin, 1998; Hair et al., 2014). A reliability coefficient estimate of 0.70 or higher (at least for ρ) is considered to be acceptable for an established construct; for emerging constructs/concepts, a lower threshold of 0.60 is acceptable (Hair et al., 2014; Nunnally & Bernstein, 1994).

Table 8.1 depicts the reliability coefficients of eight of the nine constructs of the main model. Based on the rule-of-thumb cut-off values mentioned above, it is evident that the reliability measures exceed the threshold value indicating scale reliability. It is important to note that survey questionnaire item Q38 was excluded because inclusion of Q38 reduces the composite reliability coefficient (ρ) of the construct People Satisfaction, markedly.

Table 8.1: The Reliability Statistics for Conventional Constructs

Construct	Cronbach's α	Composite reliability coefficient (ρ)	AVE
Mission	0.70	0.81	0.52
Strategy	0.85	0.90	0.69
People	0.64	0.81	0.68
Organisational Infrastructure	0.73	0.86	0.75
Financial Health			
Processes	0.92	0.92	0.67
Client Satisfaction	0.81	0.86	0.61
People Satisfaction	0.81	0.87	0.57
Donor Satisfaction	0.68	0.82	0.61

8.2.4. Construct Validity

As mentioned in Chapter Five, construct validity establishes that a construct, as operationalised by its measures, represents what it is supposed to represent (Bagozzi et al., 1991; Messick, 1995). As such, construct validity is the premier indicator of validity of the operationalisations of the constructs in a theoretical framework (Messick, 1995; Nunnally & Bernstein, 1994). In PLSBSEM, the construct validity is determined through convergent validity and discriminant validity, based on the test scores, which are observed scores of the measurement items (Chin, 1998; Fornell & Larcker, 1981; Gefen & Straub, 2005). Convergent validity refers to the relatedness of measures to their assigned construct while discriminant validity refers to the extent to which measures (indicators) that do not belong to a particular construct can be discriminated from that particular construct (Chin, 1998; Hair et al., 2014). The measure average variance extracted (AVE), which indicates the proportion of variance a construct extracts from its measures, on average (Chin, 1998), is used to examine the convergent validity of a conventional construct, along with the factor loadings. An $AVE > 0.50$ is considered to be indicative of convergent validity (Chin, 1998; Hair et al., 2014). As evidenced from Table 8.1 the AVE values of the conventional constructs exceed the threshold value, suggesting convergent validity.

The final criterion for completing the examination of the validity of a conventional construct is the discriminant validity. The researcher examined the discriminant validity using two methods: the Fornell-Larcker criterion (Fornell & Larcker, 1981) and examination of the loading/cross loading matrix. Given that square root of AVE indicates the average correlation between the measures and its assigned construct, the Fornell-Larcker criterion establishes that a construct correlates more strongly with its assigned measures than with the other constructs, implying the relatedness of the measures to their assigned construct, relative to the other constructs (Fornell & Larcker, 1981). The discriminant validity is shown when the square root of the AVE of a construct exceeds the correlations between the construct under examination and the other constructs (Fornell & Larcker, 1981). It is clear that the correlations in Table 8.2 show discriminant validity. For example, the square root of AVE of the construct Client Satisfaction is 0.78; however, the correlations between this construct and the remaining constructs in the main model are much less than 0.78.

Table 8.2: Latent variable Correlations for Discriminant Validity

	Client Satisfaction	Donor Satisfaction	Mission	Org. Infrastructure	People	People Satisfaction	Processes	Strategy
Client Satisfaction	0.78	0	0	0	0	0	0	0
Donor Satisfaction	0.59	0.78	0	0	0	0	0	0
Mission	0.63	0.69	0.72	0	0	0	0	0
Org. Infrastructure	0.58	0.53	0.43	0.86	0	0	0	0
People	0.51	0.65	0.61	0.53	0.83	0	0	0
People Satisfaction	0.57	0.55	0.61	0.56	0.66	0.76	0	0
Processes	0.69	0.74	0.67	0.64	0.75	0.72	0.82	0
Strategy	0.59	0.71	0.68	0.61	0.66	0.63	0.79	0.83

Note: The diagonal elements represent the square root of AVE of the construct, which indicates the average correlation between the construct and its measures (Chin, 1998).

The second criterion for discriminant validity looks more directly at the correlation between a measure and its assigned construct (loadings) relative to the correlations between the measure and the remaining constructs (cross-loadings). The discriminant validity of the measures is shown when their loadings exceed the cross-loadings (Gefen & Straub, 2005). Table 8.3 depicts the loadings and the cross-loadings of the measures. The loading and cross-loading pattern in Table 8.3 clearly demonstrates discriminant validity. It is important to note that Q10, Q30, Q31 and Q39 were excluded, as inclusion of these questionnaire items affects discriminant validity based on both criteria.

Table 8.3: Loadings and Cross-Loadings Based on the Reflective Model

Construct to Which the Questionnaire Item Belongs	Questionnaire Item	Construct								
		1	2	3	4	5	6	7	8	9
		People	Org. Infrastructure	Financial Health	Mission	Processes	Strategy	Donor Satisfaction	Client Satisfaction	People Satisfaction
1	Q 29	0.84	0.48	0.59	0.58	0.64	0.60	0.64	0.44	0.54
	Q 19	0.81	0.39	0.60	0.43	0.60	0.48	0.43	0.40	0.56
2	Q 34	0.48	0.90	0.57	0.34	0.61	0.61	0.43	0.48	0.51
	Q 20	0.43	0.83	0.51	0.41	0.49	0.43	0.48	0.53	0.44
3	Q 14	0.51	0.52	0.72	0.45	0.58	0.50	0.58	0.43	0.54
	Q 2	0.53	0.39	0.63	0.42	0.54	0.36	0.36	0.51	0.52
	Q 15	0.45	0.39	0.67	0.43	0.59	0.63	0.47	0.38	0.43
	Q 3	0.52	0.47	0.77	0.57	0.56	0.61	0.50	0.51	0.44
	Q 12	0.58	0.50	0.82	0.64	0.63	0.60	0.65	0.53	0.46
4	Q1	0.31	0.19	0.37	0.66	0.34	0.37	0.39	0.38	0.33
	Q5	0.31	0.26	0.42	0.70	0.43	0.41	0.43	0.40	0.29
	Q 9	0.48	0.35	0.57	0.72	0.44	0.51	0.49	0.46	0.36
	Q 13	0.58	0.39	0.61	0.79	0.64	0.62	0.62	0.55	0.69
5	Q 40	0.52	0.42	0.60	0.52	0.79	0.55	0.53	0.54	0.50
	Q 23	0.59	0.55	0.67	0.60	0.84	0.74	0.64	0.61	0.62
	Q 37	0.72	0.54	0.62	0.50	0.80	0.58	0.60	0.53	0.60
	Q 25	0.58	0.47	0.62	0.58	0.78	0.63	0.64	0.52	0.55
	Q 36	0.63	0.63	0.70	0.54	0.83	0.71	0.62	0.59	0.69
	Q 35	0.60	0.52	0.68	0.53	0.84	0.63	0.59	0.57	0.57
6	Q 16	0.48	0.52	0.62	0.57	0.61	0.85	0.57	0.53	0.52
	Q 17	0.59	0.51	0.52	0.52	0.60	0.80	0.51	0.39	0.44
	Q 24	0.56	0.44	0.68	0.63	0.70	0.80	0.67	0.54	0.53
	Q 18	0.55	0.55	0.64	0.53	0.71	0.87	0.59	0.50	0.60
7	Q 27	0.55	0.45	0.54	0.58	0.61	0.53	0.80	0.52	0.32
	Q 22	0.49	0.49	0.59	0.50	0.62	0.58	0.78	0.53	0.57
	Q11	0.47	0.25	0.54	0.54	0.47	0.54	0.76	0.29	0.38
8	Q 26	0.40	0.52	0.53	0.49	0.57	0.55	0.45	0.83	0.48
	Q 6	0.36	0.30	0.39	0.50	0.42	0.28	0.42	0.59	0.32
	Q 8	0.37	0.47	0.53	0.47	0.54	0.53	0.45	0.83	0.46
	Q 32	0.46	0.49	0.59	0.54	0.60	0.47	0.53	0.85	0.51
9	Q 21	0.53	0.45	0.51	0.47	0.56	0.49	0.46	0.31	0.72
	Q 28	0.53	0.51	0.57	0.50	0.61	0.52	0.43	0.57	0.83
	Q 41	0.43	0.27	0.44	0.51	0.52	0.45	0.42	0.38	0.71
	Q 33	0.55	0.51	0.53	0.46	0.61	0.52	0.42	0.53	0.81
	Q 7	0.43	0.32	0.38	0.36	0.40	0.37	0.33	0.34	0.70

Note: The sequence of questionnaire items was randomised to prevent systematic error (in general, consecutive questionnaire items do not capture the same construct).

8.2.4.1. Assessing the construct validity of the formative construct

As the first step of assessing the construct validity of the formative construct Financial Health, the researcher examined the convergent validity (predictive validity) of the measures of the construct using *redundancy analysis* (Chin, 1998). This was achieved by correlating the formative construct with its global measure (via PLSBSEM) by determining the standardised path coefficient (standardised structural regression coefficient) for the causal path created between the formative construct (predictor) and its global measure (response). The global measure of Financial Health was taken as the fourth questionnaire item (Q4) in part two of the questionnaire; this questionnaire item examines *projects aimed at raising funds* (Revenue Increase). The questionnaire item Q4 was treated as a global measure on Financial Health because although it is not an element that forms the meaning of Financial Health, it is nonetheless expected to be correlated markedly with Financial Strength, because most NPOs engage in fundraising projects to boost their revenue. The rule of thumb cut-off value for the correlation between a formative construct and its global measure is 0.8 to confirm the convergent validity (Hair et al., 2014). The redundancy analysis showed that the standardised path coefficient (hence the correlation) between the construct and its global measure is 0.853 ($p < 0.001$), which is well above the cut-off value (0.8) prescribed by Hair et al., thus confirming the convergent validity of Financial Health.

As the second step of examining construct validity of the formative construct, the researcher examined the collinearity of the measures of Financial Health (Q14, Q15, Q3, Q12, Q2); since the score of the Financial Health (response) is obtained by regressing Financial Health with its five predictors (the measures that form the construct), collinearity examination of the measures becomes relevant (Hair et al., 2014). The statistical software package IBM SPSS 22 was used to examine collinearity. The collinearity statistics, namely *the variation inflation factor* (VIF) values of the five measures of Financial Health were found to be 1.25, 1.53, 1.79, 1.58, and 1.51 for Q14, Q15, Q3, Q12, and Q2 respectively. This indicated that there is no significant collinearity (because $VIF < 5.0$) involving the measures of the formative construct.

The final step of scrutinising the validity of Financial Health involved examining the statistical significance and relevance of the weights of the measures. Since PLSBSEM

is a nonparametric approach (Chin, 1998), as mentioned earlier, the statistical significance of all model parameters is examined using a re-sampling method; SmartPLS2 uses the resampling technique “bootstrapping” (Efron, 1979) to generate the t values of the parameters for statistical significance (Hair et al., 2014). The statistical significance of the weights was examined using the p values of the t statistics. Based on the p values (0.000, 0.001, 0.000, 0.000, and 0.001 for Q14, Q15, Q3, Q12, and Q2 respectively), all measures were deemed significant and relevant at 5% level.

Now, having established the validity of all the constructs in the main model, the researcher completes the validation of the PM dimensions (constructs) in her PM model, including their operational definitions.

8.3. TESTING THE HYPOTHESISED THEORETICAL RELATIONSHIPS

A hypothesised theoretical relationship is supported by data, when it can be shown that the relationship is statistically significant, given the specified level of significance, which is 5% (statistical conclusion validity). In a PLSBSEM sense, this equates to examining the statistical significance of the estimated structural regression coefficients of the hypothesised relationships. In addition to statistical significance (a zero relationship versus a nonzero relationship), examining the practical significance (the magnitude of the relationship) is also important in quantitative social and behavioural research (Bryman, 2012; Chin, 1998; Hair et al., 2014). In PLSBSEM, the practical significance of a hypothesised relationship/s is examined by scrutinising the magnitude of the estimated structural relationship/s (Chin, 1998; Hair et al., 2014). Therefore, the estimated standardised regression coefficients of the hypothesised structural relationships and the amount (proportion) of variability explained by the independent latent variables (i.e. the R^2 of the dependent latent variables) become important. Figure 8.1 depicts the structural model, along with the following model parameters estimated by the PLSBSEM algorithm:

- The standardised structural regression coefficients of the hypothesised relationships and their statistical significance;
- The R^2 values of the dependent latent variables (constructs).

The prerequisite for validating the hypothesised relationships in structural equation modelling (SEM) approaches is an acceptable level of goodness of fit of the model to data, overall (i.e. taking into consideration all the specified statistical relationships, as a whole). Therefore the *overall goodness of fit* of the model, also known as the *global goodness of fit*, is examined next.

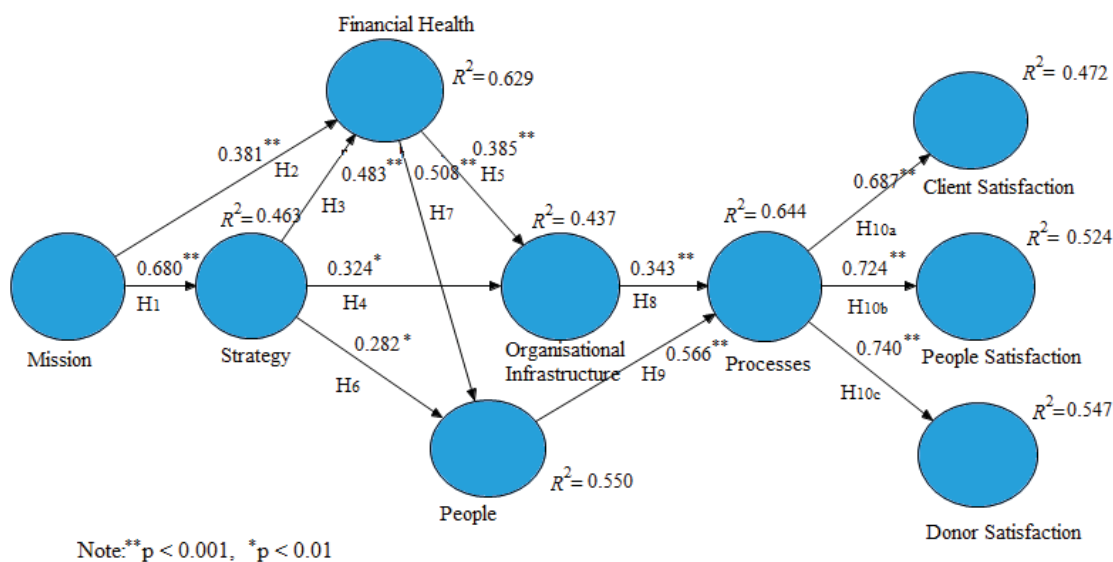


Figure 8.1: Estimated structural model parameters and the significance of hypothesised relationships

8.3.1. The Overall Goodness of Fit of the Model

As mentioned in Chapter Five, in any SEM approach, the researcher specifies two types of statistical relationships: the relationships the measures of the constructs have with their assigned constructs (this is known as the measurement model, which is also known as the outer model in PLSBSEM) and the hypothesised relationships between the constructs themselves (this is known as the structural model, which is also known as the inner model in PLSBSEM). As mentioned above, and also in Chapter Five, in SEM, the overall goodness of fit refers to a single metric (or parameter) that examines the goodness of fit of all the specified statistical relationships to the data, as a whole. Unfortunately, as mentioned earlier, there is no well-established global (single) goodness of fit measure to conclude whether or not the model as a whole (the measurement and structural models combined), is a good fit to data (Chin, 1998; Grigg & Jayamaha, 2014; Hair et al., 2014). Thus in PLSBSEM, as mentioned in Chapter Five, the quality of the overall model is assessed through assessment of the quality of

the measurement model and the structural model separately (Chin, 1998; Hair et al., 2014). In PLSBSEM, communalities (or loadings), composite reliability (ρ) and AVE are used to test the goodness of the measurement model⁴⁵, while the *coefficient of determination* (R^2), *predictive relevance* (Q^2), *size and significance of the path coefficients* (*standardised structural regression coefficients*), f^2 *effect sizes*, and q^2 *effect sizes* are used to examine the goodness of the structural model (Chin, 1998; Hair et al., 2014). The researcher observes that out of the above many different useful indicators for assessing the structural model, only the R^2 and size/significance of the path coefficients are reported and discussed in journal articles.

Since the structural model consists of a series of multiple regression equations, before proceeding further, the researcher checked the structural model for collinearity issues. The VIFs of all predictor constructs (Table 8.4) are below the threshold value of 5.0 for the predictor constructs to be considered as collinear (Hair et al., 2014). Therefore, the researcher argues that although there is strong correlation between the constructs, they do not threaten model adequacy due to any collinearity. Moreover, the researcher observes that past studies that examine the theoretical validity of the PM models (e.g. Jayamaha et al., 2009; Saunders, Mann, & Grigg, 2008) report similar strong relationships between constructs.

Table 8.4: Collinearity Assessment

First set (Predictors of Financial Health)		Second Set (Predictors of People)		Third Set (Predictors of Organisational Infrastructure)		Fourth Set (Predictors of Processes)	
Constructs	VIF	Constructs	VIF	Constructs	VIF	Constructs	VIF
Mission	1.89	Strategy	2.21	Financial Health	2.21	People	1.39
Strategy	1.89	Financial Health	2.21	Strategy	2.21	Organisational Infrastructure	1.39

Since the R^2 values associated with multiple regression depend on the research discipline and model complexity, there's no rule of thumb for acceptable R^2 values; a higher R^2 value indicates that much of the variability of the dependent variable is

⁴⁵ The goodness of the measurement model was found to be satisfactory (these were examined in section 8.2).

explained by the independent variables (predictors). In social sciences, a R^2 value above 0.5 is considered to be a markedly strong relationship, and hence acceptable, from a practical perspective (Hair et al., 2014). It becomes clear that all the R^2 values (Figure 8.1) in the model thus become acceptable; also all the hypothesised relationships appear significant both statistically ($p \ll 0.05$) and practically (the estimated path coefficients are significantly greater than zero).

As the next step, although being regarded as an optional test, the researcher examined the predictive relevance of the structural model using the ‘blindfolding procedure’ incorporated in SmartPLS software. This blindfolding procedure estimates a predictive relevance measure for *each endogenous*⁴⁶ *reflective construct*⁴⁷, known as the Stone and Geisser’s Q^2 (Geisser, 1974; Stone, 1974). A brief description of Stone and Geisser’s blindfolding algorithm is as follows (for details see Hair et al., 2014). Step 1: the software would omit data in the every seventh observation in the data matrix (this omission distance can be changed but the default is every seventh observation traversing from left to right); Step 2: the software would impute this missing data using the specified model to predict the blindfolded data; Step 3: the software would then determine the error (actual values minus the imputed value) for each observation to compute an error sum of squares; Q^2 would then be calculated using the formula $Q^2 = 1 - (\text{error sum of squares associated with the omitted values})/(\text{observed sum of squares of the omitted values})$. The necessary and sufficient condition for predictive relevance is $Q^2 > 0$ (Chin, 1998; Hair et al., 2014). This procedure does not apply for formative endogenous constructs because the scores of a measure of a formative construct cannot be predicted from the score of the formative construct because the score of a construct cannot predict the score of a measure (Chin, 1998). The Q^2 values (Table 8.5) of all the endogenous reflective constructs are above zero, thus confirming predictive relevance on the endogenous constructs.

⁴⁶ As explained in Chapter 5, an endogenous construct is a construct that is being predicted (explained) by other constructs (i.e. predictor construct).

⁴⁷ As explained in Chapter 5, a reflective construct is a construct that is modelled using conventional wisdom; in the positivistic epistemology, which the researcher refers to as conventional wisdom, a construct is believed to exist irrespective of the observer and that the measures are mere reflections (manifestations) of the construct.

Table 8.5: Predictive Relevance of Endogenous Reflective Constructs

Construct	Q^2 value
Strategy	0.308
People	0.373
Processes	0.447
Client satisfaction	0.239
People Satisfaction	0.240
Donor Satisfaction	0.290
Organisational Infrastructure	0.255

Finally, the f^2 and q^2 effect sizes were examined for endogenous constructs that are being predicted (explained) by multiple (two or more) constructs. In the researcher's hypothesised theoretical model (Figure 8.1), there are four such constructs: Processes, People, Organisational Infrastructure, and Financial Health. The effect size f^2 in relation to a predictor construct (independent variable) indicates the size of effect that predictor construct has in predicting the endogenous construct, in terms of changing the R^2 of the endogenous construct (details in Chapter Five). By the same token, effect size q^2 indicates the size of effect that the predictor construct has in predicting the endogenous construct, in terms of changing the Q^2 of the endogenous construct. In terms of Cohen's original classification of effect sizes (Cohen, 1988), f^2 and q^2 values 0.02, 0.15, and 0.35 are taken as small, medium and large effect sizes respectively, for the predictor construct under review. Table 8.6 depicts the effect sizes of the predictor constructs that predict the four endogenous constructs mentioned earlier. It is evident from the Table 8.6 that most of the f^2 and q^2 effect sizes are close to (or exceed) a medium effect size (= 0.15) and the ones that have smaller effect sizes are larger than the 0.02 value, which is being regarded as the small effect size. Thus overall, the f^2 and q^2 effect sizes shown in Table 8.6 indicate that the predictors (independent variables) associated with the hypothesised relationships are relevant (have a sizable effect), in a statistical sense.

Table 8.6: The Effect Sizes of the Predictors of the Relevant Endogenous Constructs

	Response: Organisational Infrastructure		Response: People	
Predictor	f^2 Effect Size	q^2 Effect Size	f^2 Effect Size	q^2 Effect Size
Financial Health	0.101	0.069	0.148	0.073
Strategy	0.085	0.057	0.102	0.056
	Response: Processes			
Predictor	f^2 Effect Size	q^2 Effect Size		
People	0.579	0.277		
Organisational Infrastructure	0.214	0.145		
	Response: Financial Health			
Predictor	f^2 Effect Size	q^2 Effect Size		
Mission	<i>Not applicable, being a formative construct</i>			
Strategy				

As mentioned earlier, based on the parameter estimates (and their p values where relevant) shown in Figure 8.1, all hypothesised relationships become statistically and practically significant, based on the p values ($\ll 0.05$) and the sizes of the path coefficients (also the R^2 values) respectively. The fact that all path coefficients (standardised structural regression coefficients) of the structural model being significant, means that all ten hypotheses specified by the researcher (Figure 8.1) have been supported by the data (the tenth hypothesis H_{10} has three subparts and as shown in Figure 8.1, all three of them have been well supported by the data). Thus theoretical validation of the PM model was completed within an Australasian nonprofit healthcare context.

As regards the validity of the operational definitions of the 9 constructs of the model, as mentioned earlier, 5 out of the 41 survey items (12%) were discarded on statistical grounds. The researcher does not consider this as a serious issue, in the context of the study. The researcher derived the operational definitions (the survey items in the questionnaire) from a small sample study (qualitative data collection) and tried to generalise these operational definitions across her sampling frame. Under this

circumstance, it is not surprising that 12% of the survey items had to be deleted. However, it is interesting to conduct more fieldwork to examine why certain key measurement aspects are not generalisable across healthcare NPOs. For example, Q30 (staff training and creating a high performance work culture) was not supported by data, as a performance measure of People. It could be quite possible that healthcare NPOs that rely heavily on volunteer effort do not place human resource development high on their agenda. Identifying why people work in NPOs as volunteers and what is required for them to improve their performance, is beyond the scope of this study.

8.4. DISCUSSING THE RESULTS OF THE MAIN MODEL FROM A THEORETICAL PERSPECTIVE

One of the most important outcomes of the study is adding clarity to the theory underlying the BSC (the nonprofit version) through deductive reasoning. The BSC has been criticised for lack of a unified theory (Anon, 2008; Kong, 2008; Nørreklit, 2000; Williams & Lewis, 2008). The inventors of the BSC, Robert Kaplan and David Norton asserted that causality, as far as the BSC is concerned, comes from the specific strategy map (the linkages between the strategic objectives) of the organisation (Kaplan, 2008). As the researcher found in her fieldwork (stage 1 data collection), strategic objectives vary from organisation to organisation (also, some organisations were found to have few strategic objectives while others were found to have several). While strategy maps would be useful for individual organisations (as long as managers are able to develop them) because these cause managers to view their organisation as a system, strategy maps provide little insights for the generalisability of the BSC. In fairness to Kaplan and Norton, they did acknowledge the theoretical and practical issues involving strategy maps; as a remedy for these issues, Kaplan suggested that “one might even describe the generic strategy map as a dumbed-down representation of causal linkages” to explain the theoretical basis of the BSC (Kaplan, 2008, p. 1268). Also, in articulating the nonprofit version of the BSC, Kaplan (2001) asserted that the strategy should be quantified and measured. Similarly, he contended that the Mission of a NPO should be “featured and measured” (p. 361); however, the BSC provides little guidance as to how this could be accomplished. The researcher addressed these gaps in this study.

By successfully examining the theoretical validity of the BSC to the nonprofit healthcare sector, the researcher showed how to achieve the organisational Mission, which results in satisfying the key stakeholders. The results validated the underlying theory (presented as a series of hypotheses as depicted in Figure 8.1) that Stakeholder Satisfaction results from managing the Processes efficiently and effectively by using the organisational resources—specifically, the Financial Health, People, and Knowledge Infrastructure, which are being driven by the organisational Strategy.

This study also brings clarity to the nonprofit BSC, because the researcher explains how organisational Mission drives the Strategy (which in turn drives the system) to achieve the stakeholder (Client, People, and the Donors) outcomes. Although the Mission has been used in the nonprofit BSC to mean both the driver and the final outcome, the nonprofit BSC fails to answer what precisely the Mission represents (unless a strategy map is juxtaposed upon the PM dimensions). The researcher has filled this gap through her model, which is able to develop performance measures for the Mission itself as well as the Strategy (details were given in Chapter Six). Measuring the Mission and Strategy helps an organisation to improve the actual driving elements of performance improvement, which in turn increases the ‘board accountability’ to the key stakeholders.

The greatest contribution the researcher makes to the existing body of knowledge, in empirically validating her theoretical model, is that the researcher’s theoretical model can be used as a tool to explain successful performance improvement (intervention) as well as unsuccessful performance improvement. This is because the researcher’s theoretical model explains how a NPO achieves Stakeholder Satisfaction, and thereby the Mission. The research findings became even more valuable because it was found that the explanatory constructs have high predictive relevance (explanatory power). Another key theoretical contribution of the study is operationalisation of PM dimensions. These operationalisations can be used by NPOs for self-assessment (e.g. the quantitative questionnaire used in this study) or to design organisation specific performance measures for each PM dimension (details in Chapter Six).

8.5. DISCUSSING THE RESULTS ON THE MAIN MODEL FROM A PRACTICAL PERSPECTIVE

The findings confirmed that the Mission drives both the Strategy (0.680) and the Financial Health (0.381); the figures in parenthesis refer to the estimates of the standardised structural regression coefficients. This implies the importance of the Mission in strategic planning. The Mission is the starting point in strategic planning (Bryson, 2011). As such, managers need to ensure that the Mission is subsumed in the strategies being formulated. Further, the model implies that Strategy is fundamental to organisational performance in that it is important to formulate the right strategies to ensure that People are focused on delivering the Strategy. It is reasonable to assume that the direct relationship between the Mission and the Financial Health would be weaker in organisations that rely entirely on government funding, because the Mission → Financial Health link, as argued earlier (section 6.4.4), represent private donations (philanthropy). This leads to an interesting consideration. There could be moderators that affect the hypothesised causal relationships, which should be studied separately. For example, for the link described above, the level of nongovernmental income (donations and fees) could affect to what extent the Mission has a direct effect on the Financial Health of a NPO.

The findings also suggest that the Strategy has a sizable positive effect on the following: (a) Organisational Infrastructure (0.324), (b) People (0.282), and (c) the organisation's Financial Health (0.483). Of these three relationships, the latter being the strongest as suggested by the standardised structural regression coefficients shown in parenthesis. Thus, consolidating the financial position seems to be one of the key direct outcomes of an effective Strategy of a healthcare NPO. Since the Organisational Infrastructure as well as People include the strengths and weaknesses of the NPO's capital—employees, technology, skills, knowledge—that enable the NPO to carry out its processes efficiently and effectively, it is justifiable that the Strategy has a sizable positive effect on the said constructs (PM dimensions). This is because one role of the Strategy would be to enhance the capital. Through the preceding fieldwork (the qualitative case study), it was found that Australasian healthcare NPOs use their tacit knowledge (the knowledge that the organisations assimilate over the course of its existence) to do things better (i.e. improve processes). Human resources (People) are also a key resource for

the NPOs (Laratta, 2010; Salamon & Anheier, 1997; Srikanthan & Dalrymple, 2003) and the preceding fieldwork also found that NPOs do compete with the for-profit organisations in attracting certain skilled people, hence the need to have an effective strategy to attract such people.

In finding out how the Financial Health (which includes the cash inflows as well as cash outflows, such as direct costs and overheads) works as an enabler in achieving Stakeholder Satisfaction through Processes, it becomes clear from the estimates of the standardised structural regression coefficients (Figure 8.1) that Financial Health has a greater effect on Processes through People than through Organisational Infrastructure. Needless to say, although NPOs are not profit oriented organisations, they nonetheless have to have the financial infrastructure to continue serving their customers through its People, by way of paying the employees (non-volunteers), meeting numerous overheads, and enhancing organisational learning, through both People and the Organisational Infrastructure (physical and information infrastructure).

In the model, in keeping with the systems perspective, the Processes act as a solitary endogenous variable that predicts and explains Stakeholder Satisfaction (and thereby the Mission). It becomes evident from the estimates of standardised structural regression coefficients and the R^2 values (Figure 8.1) that satisfaction of all three stakeholder groups (Clients, People, and Donors) are equally (approximately) strongly explained by the Processes. This also seems to be consistent with the *normative stakeholder theory*, which implies that no legitimate stakeholder should receive more priority than the other.

As regards process inputs, it seems from the standardised structural regression coefficients that People have a stronger effect (0.566) on the Processes than the Organisational Infrastructure does (0.343), which may reflect the intangible nature of the service processes of healthcare NPOs. However, since Stakeholder Satisfaction was operationalised in part by asking the respondents about the Processes that the NPO has put in place to measure stakeholder satisfaction, it is possible that organisations that do well in Processes also (to some extent) show up as doing well in satisfying the stakeholders (and vice versa). This is an inherent limitation in a self-administered survey. In designing the survey questionnaire, the researcher decided that it would be more accurate to include some survey items that *indirectly* examine Stakeholder

Satisfaction; the researcher sought responses for questions about Processes NPOs put in place to measure Stakeholder Satisfaction, rather than relying solely on direct questions on Stakeholder Satisfaction. In surveys, it is typical to include indirect questions (proxy measures) instead of direct measures to obtain truer and more open responses from the respondents (Fowler, 2013; Nardi, 2015).

The model and its associated parameter estimates (Figure 8.1) show to what extent the Mission drives the system (through Strategy and Financial Health) to achieve mission-related outcomes: Client Satisfaction, People Satisfaction, and Donor Satisfaction. The Mission of any NPO is to create value for society (or any target group), by being both efficient and effective in managing its Processes (McDonald, 2007). In fact, for NPOs, Mission is the main reason for their existence (Kaplan, 2001). The model (more technically, the measurement model) also operationalises the Mission; being able to measure Mission is a significant practical contribution of the study. The *social impact* is frequently used as a means to determine the extent to which a NPO creates value to society (Dees, Anderson, & Wei-Skillern, 2004; Kaplan, 2001; Richmond, Mook, & Jack, 2003). Clark, Rosenzweig, Long, and Olsen (2004) define social impact (social return/social value creation) as “*the portion of the total outcome that happened as a result of the activity of an organisation, above and beyond what would have happened anyway*” (emphasis added). The researcher’s model, which takes a systems/process perspective, agrees with this notion. This is because the researcher incorporated operational descriptions for social impact (social profit) creation, under Mission (see Chapter Six). The researcher observed that the corporate sector also put social impact high on the agenda. For example, Emerson (2003) found that a number of corporate CEOs emphasise the positive social impact created by their corporations, as a strategy to increase the value of the corporation.

Since the researcher operationalised the constructs (PM dimensions) of her model (all but the five questionnaire items were found to be valid), as mentioned earlier, these operational descriptions/definitions can readily be used by a NPO to generate its own performance measures (details were provided in Chapter Six). Unlike in the researcher’s model, Mission and Strategy are not measured in the BSC directly. The BSC measures performance further down the stream. The researcher believes that by

being able to directly measure the Mission and the Strategy, NPOs can react to strategy ineffectiveness more swiftly than otherwise.

8.5.1. Concluding Discussion on the Third Research Objective

In section 8.3.1, the overall goodness of fit of the model to the data was examined using methods applicable to PLSBSEM; results indicated that the model, as a whole, fits the data satisfactorily. Moreover, all of the hypothesised relationships in the structural model were statistically significant and the dependent latent variables were well explained by the independent (predictor) latent variables, in terms of R^2 . The third research question (RQ3) asks: *“How do the PM dimensions theoretically relate to one another in explaining the achievement of strategic objectives of a NPO?”* Through the conceptual model and subsequent fieldwork (the case study involving qualitative data), the researcher explained in the form of hypotheses how her PM dimensions theoretically relate to one another in explaining the achievement of strategic objectives of a NPO. In this chapter, the researcher showed that her model fits the data well and that all her hypothesised relationships stay in contention (they were all statistically significant, based on the p values). Therefore, through the results reported in this chapter, the researcher reinforced her explanations of the proposed theoretical relationships between PM dimensions to answer RQ3.

In addition, the results of the main model were discussed from theoretical and practical perspectives in sections 8.4 and 8.5 respectively. With this, the researcher concludes her discussion on the third research objective: to test the theoretical causal relationships (i.e. the structural relationships) between BSC measurement domains and interpret the results from a theoretical and practical perspective.

8.6. ANALYSING THE COUNTRY EFFECT ON THE HYPOTHESISED RELATIONSHIPS

To people outside Australia and New Zealand, there is hardly any difference between the two Tasman neighbours, in regard to people and social entities. However, researchers (e.g. Avery, Everett, Finkelde, & Wallace, 1999; Felzensztein, Stringer, Benson-Rea, & Freeman, 2014; Gandolfi, 2005; Trevor-Roberts, Ashkanasy, & Kennedy, 2003) have attempted to identify differences between Australia and New

Zealand on such factors as organisations, management practices, and behavioural attributes of the leaders and their subordinates. Avery et al. (1999) show that New Zealand leaders give more autonomy to their followers (subordinates) than their Australian counterparts. Trevor-Roberts et al. (2003) hypothesised that although both Australia and New Zealand are egalitarian societies, Australian leaders are more assertive, individualistic, egotistical (confident and arrogant that they have done things right), and adventurous (care less about rules) than their New Zealand counterparts. Based on their survey data, the *t* tests for group differences were found to be small, but statistically significant at 5% level. Gandolfi (2005) found that New Zealand senior managers are more employee-centred than their Australian counterparts (albeit in downsizing decisions). The generalisability of this study, however, is questionable because the details of the case study (e.g. the number of participants, case study protocol) were not reported.

Although descriptive statistics on the researcher's survey data (Chapter Seven) did not show significant mean differences between the two countries for most survey items, in this section, the researcher examines to what extent country effect seems to moderate (changes) the hypothesised relationships of her theoretical model (Figure 8.1), on the strength of extant literature described above. However, because there is no strong evidence in the literature to formulate *a priori* hypotheses on the nature of the country effect for each relationship, the following twelve null hypotheses were posited.

Hc₁: Mission has an equal positive effect on Strategy in Australia as in New Zealand.

Hc₂: Mission has an equal positive effect on Financial Health in Australia as in New Zealand.

Hc₃: Strategy has an equal positive effect on Financial Health in Australia as in New Zealand.

Hc₄: Strategy has an equal positive effect on Organisational Infrastructure in Australia as in New Zealand.

Hc₅: Financial Health has an equal positive effect on Organisational Infrastructure in Australia as in New Zealand.

Hc₆: Strategy has an equal positive effect on People in Australia as in New Zealand.

Hc₇: Financial Health has an equal positive effect on People in Australia as in New Zealand.

Hc₈: *Organisational Infrastructure has an equal positive effect on Processes in Australia as in New Zealand.*

Hc₉: *People have an equal positive effect on Processes in Australia as in New Zealand.*

Hc_{10a}: *Processes have an equal positive effect on Client Satisfaction in Australia as in New Zealand.*

Hc_{10b}: *Processes have an equal positive effect on People Satisfaction in Australia as in New Zealand.*

Hc_{10c}: *Processes have an equal positive effect on Donor Satisfaction in Australia as in New Zealand.*

The above hypotheses were tested using the same data set ($n = 223$) used to test the main theoretical model. The Australian data ($n = 117$) and New Zealand data ($n = 106$) were fitted to the main theoretical model separately, thus creating two statistical models, to derive two sets of parameter estimates for each country. For ease of reference, these two statistical models were named the Australian model (Figure 8.2) and the New Zealand Model (Figure 8.3).

8.6.1. Results on the Australian Model

Prior to interpreting the parameter estimates (Figure 8.2), the researcher examined scale reliability (Table 8.7) and construct validity (Table 8.8) information using the same approach that was used earlier. The results in these two tables demonstrate reliability and validity of the constructs that constitute the hypotheses, based on Australian data.

Table 8.7: Reliability Statistics of the Australian Model

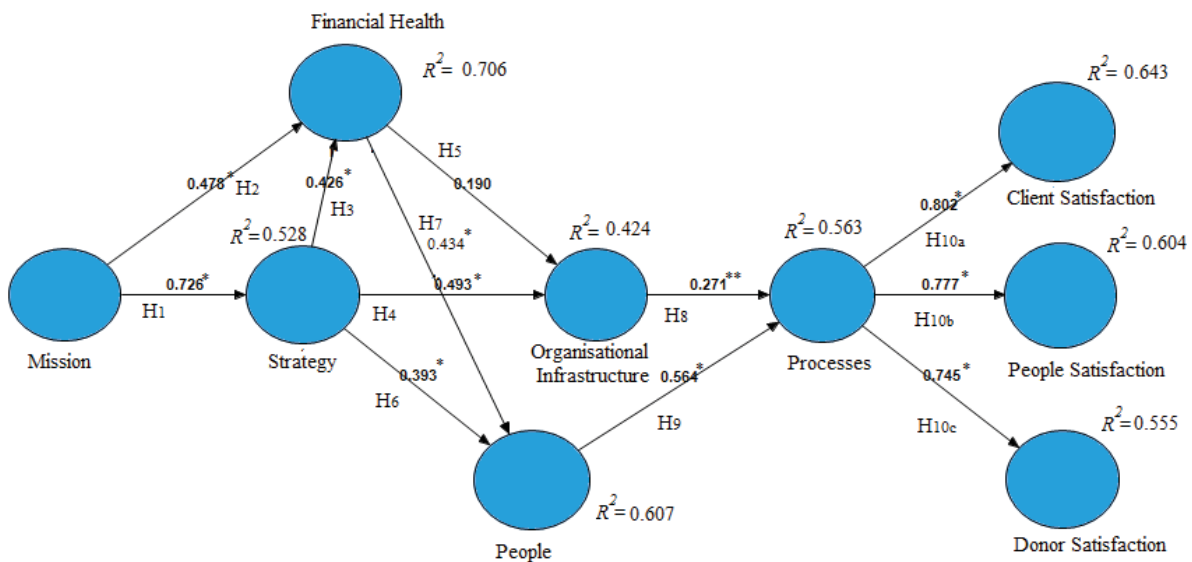
Construct	Cronbach's α	Composite reliability coefficient (ρ)	AVE
Mission	0.69	0.81	0.52
Strategy	0.87	0.91	0.72
People	0.62	0.84	0.73
Organisational Infrastructure	0.78	0.90	0.82
Financial Health			
Processes	0.93	0.94	0.73
Client	0.83	0.90	0.67

Construct	Cronbach's α	Composite reliability coefficient (ρ)	AVE
Satisfaction			
People Satisfaction	0.80	0.86	0.56
Donor Satisfaction	0.71	0.83	0.62

Table 8.8: Loading and Cross-Loading Information for the Australian Model to Examine Construct Validity

Construct to Which the Questionnaire Item Belongs	Questionnaire Item	Construct								
		1	2	3	4	5	6	7	8	9
		People	Org. Infrastructure	Financial Health	Mission	Processes	Strategy	Donor Satisfaction	Client Satisfaction	People Satisfaction
1	Q 29	0.88	0.61	0.62	0.56	0.59	0.78	0.64	0.39	0.56
	Q 19	0.83	0.60	0.52	0.42	0.57	0.59	0.54	0.49	0.54
2	Q 34	0.49	0.85	0.64	0.53	0.64	0.45	0.42	0.49	0.68
	Q 20	0.69	0.77	0.51	0.47	0.49	0.57	0.52	0.44	0.40
3	Q 14	0.43	0.48	0.62	0.47	0.60	0.41	0.60	0.44	0.51
	Q 2	0.37	0.56	0.62	0.52	0.60	0.33	0.40	0.59	0.53
	Q 15	0.53	0.55	0.79	0.59	0.65	0.67	0.57	0.45	0.56
	Q 3	0.55	0.55	0.81	0.63	0.52	0.65	0.51	0.42	0.40
	Q 12	0.54	0.50	0.77	0.65	0.63	0.61	0.59	0.57	0.48
4	Q1	0.37	0.36	0.56	0.73	0.49	0.47	0.55	0.47	0.32
	Q5	0.34	0.39	0.45	0.71	0.40	0.48	0.47	0.48	0.24
	Q 9	0.46	0.45	0.57	0.67	0.43	0.55	0.54	0.45	0.31
	Q 13	0.49	0.55	0.66	0.78	0.64	0.58	0.59	0.60	0.70
5	Q 40	0.53	0.58	0.68	0.59	0.89	0.60	0.58	0.68	0.65
	Q 23	0.62	0.54	0.72	0.66	0.85	0.76	0.70	0.72	0.63
	Q 37	0.56	0.68	0.69	0.56	0.83	0.56	0.57	0.67	0.71
	Q 25	0.55	0.57	0.67	0.65	0.84	0.64	0.69	0.71	0.60
	Q 36	0.65	0.61	0.63	0.51	0.84	0.66	0.66	0.66	0.72
	Q 35	0.56	0.59	0.66	0.56	0.88	0.65	0.62	0.66	0.67
6	Q 16	0.68	0.55	0.73	0.72	0.69	0.89	0.70	0.57	0.55
	Q 17	0.78	0.57	0.57	0.49	0.51	0.81	0.55	0.39	0.41

Construct to Which the Questionnaire Item Belongs	Questionnaire Item	Construct								
		1	2	3	4	5	6	7	8	9
		People	Org. Infrastructure	Financial Health	Mission	Processes	Strategy	Donor Satisfaction	Client Satisfaction	People Satisfaction
	Q 24	0.54	0.46	0.67	0.67	0.73	0.79	0.71	0.59	0.52
	Q 18	0.75	0.50	0.63	0.56	0.62	0.89	0.66	0.48	0.54
7	Q 27	0.66	0.55	0.61	0.72	0.64	0.73	0.87	0.60	0.45
	Q 22	0.49	0.42	0.58	0.55	0.70	0.54	0.83	0.69	0.60
	Q11	0.53	0.38	0.51	0.52	0.33	0.59	0.65	0.26	0.33
8	Q 26	0.41	0.43	0.51	0.54	0.71	0.51	0.63	0.87	0.49
	Q 6	0.24	0.35	0.31	0.47	0.40	0.25	0.38	0.58	0.38
	Q 8	0.47	0.49	0.61	0.66	0.71	0.57	0.61	0.88	0.55
	Q 32	0.51	0.58	0.63	0.62	0.75	0.57	0.65	0.91	0.63
9	Q 21	0.57	0.58	0.53	0.47	0.51	0.52	0.49	0.34	0.69
	Q 28	0.48	0.61	0.59	0.51	0.68	0.46	0.57	0.62	0.86
	Q 41	0.37	0.41	0.43	0.40	0.60	0.40	0.40	0.40	0.68
	Q 33	0.61	0.53	0.53	0.43	0.67	0.55	0.49	0.61	0.81
	Q 7	0.31	0.35	0.26	0.23	0.37	0.22	0.24	0.30	0.67



Note: *p < 0.001, **p < 0.05

Figure 8.2: Parameter estimates in the Australian model

As shown in Figure 8.2, all path coefficients (standardised structural regression coefficients) in the Australian model (Figure 8.2) but one (the path relationship:

Financial Health → Organisational Infrastructure) appear significant. The reason for appearance of a nonsignificant path was assumed to be the smaller sample size ($n = 117$). In any case, the R^2 values associated with the endogenous constructs indicate the validity of the hypothesised structural relationships in the model.

8.6.2. Results on the New Zealand Model

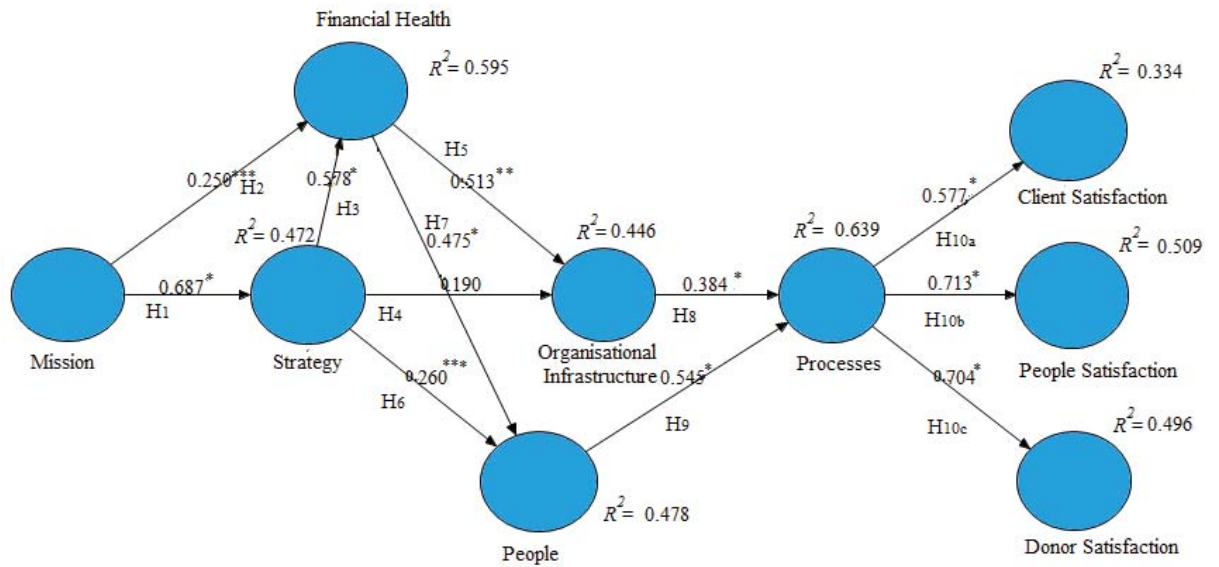
The results in Table 8.9 and Table 8.10 demonstrate reliability and validity of the constructs based on New Zealand data.

Table 8.9: Reliability Statistics of the New Zealand Model

Construct	Cronbach's α	Composite reliability coefficient (ρ)	AVE
Mission	0.64	0.77	0.47
Strategy	0.84	0.89	0.68
People	0.48	0.79	0.66
Organisational Infrastructure	0.69	0.86	0.75
Financial Health			
Processes	0.91	0.93	0.69
Client Satisfaction	0.82	0.88	0.66
People Satisfaction	0.82	0.87	0.58
Donor Satisfaction	0.63	0.80	0.58

Table 8.10: Loading and Cross-Loading Information for the New Zealand Model to Examine Construct Validity

Construct to Which the Questionnaire Item Belongs	Questionnaire Item	Construct								
		1	2	3	4	5	6	7	8	9
		People	Org. Infrastructure	Financial Health	Mission	Processes	Strategy	Donor Satisfaction	Client Satisfaction	People Satisfaction
1	Q 29	0.81	0.42	0.49	0.50	0.60	0.56	0.52	0.35	0.54
	Q 19	0.81	0.33	0.59	0.26	0.57	0.44	0.44	0.32	0.39
2	Q 34	0.50	0.92	0.62	0.38	0.66	0.62	0.51	0.46	0.49
	Q 20	0.27	0.82	0.51	0.37	0.41	0.33	0.39	0.48	0.49
3	Q 14	0.61	0.55	0.84	0.51	0.69	0.60	0.66	0.47	0.56
	Q 2	0.51	0.39	0.60	0.28	0.54	0.38	0.23	0.42	0.50
	Q 15	0.37	0.41	0.64	0.39	0.62	0.57	0.49	0.43	0.37
	Q 3	0.43	0.48	0.72	0.58	0.52	0.56	0.54	0.60	0.54
	Q 12	0.42	0.51	0.77	0.59	0.61	0.65	0.57	0.52	0.53
4	Q1	0.08	0.15	0.19	0.54	0.19	0.25	0.21	0.36	0.35
	Q5	0.17	0.24	0.32	0.69	0.42	0.40	0.39	0.30	0.38
	Q 9	0.40	0.38	0.51	0.70	0.39	0.49	0.40	0.30	0.37
	Q 13	0.47	0.34	0.59	0.77	0.61	0.62	0.56	0.53	0.70
5	Q 40	0.56	0.49	0.73	0.58	0.86	0.64	0.62	0.51	0.59
	Q 23	0.59	0.60	0.72	0.57	0.89	0.74	0.62	0.56	0.67
	Q 37	0.70	0.56	0.64	0.46	0.76	0.63	0.55	0.44	0.58
	Q 25	0.47	0.44	0.53	0.46	0.74	0.63	0.55	0.38	0.49
	Q 36	0.67	0.54	0.73	0.62	0.86	0.76	0.59	0.53	0.69
	Q 35	0.61	0.54	0.72	0.48	0.88	0.68	0.59	0.45	0.52
6	Q 16	0.41	0.52	0.59	0.54	0.61	0.83	0.64	0.47	0.45
	Q 17	0.53	0.39	0.48	0.53	0.61	0.80	0.54	0.29	0.44
	Q 24	0.48	0.48	0.69	0.58	0.66	0.83	0.58	0.50	0.50
	Q 18	0.60	0.49	0.68	0.60	0.80	0.83	0.64	0.45	0.62
7	Q 27	0.53	0.40	0.49	0.48	0.57	0.55	0.77	0.42	0.33
	Q 22	0.40	0.53	0.60	0.41	0.56	0.58	0.73	0.43	0.52
	Q11	0.40	0.23	0.45	0.51	0.45	0.52	0.78	0.24	0.34
8	Q 26	0.47	0.52	0.64	0.53	0.57	0.52	0.47	0.90	0.52
	Q 6	0.22	0.30	0.32	0.45	0.26	0.28	0.32	0.53	0.26
	Q 8	0.28	0.44	0.55	0.42	0.46	0.42	0.43	0.87	0.49
	Q 32	0.32	0.46	0.60	0.44	0.52	0.45	0.38	0.90	0.52
9	Q 21	0.48	0.41	0.47	0.41	0.52	0.45	0.35	0.21	0.70
	Q 28	0.50	0.48	0.63	0.59	0.66	0.51	0.39	0.54	0.83
	Q 41	0.34	0.30	0.44	0.61	0.46	0.47	0.47	0.42	0.71
	Q 33	0.40	0.49	0.58	0.48	0.57	0.45	0.40	0.50	0.82
	Q 7	0.47	0.42	0.54	0.56	0.49	0.48	0.44	0.48	0.75



Note: *p < 0.001, **p < 0.01, ***p < 0.05

Figure 8.3: Parameter estimates in the New Zealand model

As shown in Figure 8.3, all path coefficients (standardised structural regression coefficients) in the New Zealand model (Figure 8.3) but one (the path relationship: Strategy → Organisational Infrastructure) appear significant. Again, the reason for appearance of a nonsignificant path was assumed to be the smaller sample size ($n = 106$). In any case, the R^2 values associated with the endogenous constructs indicate the validity of the hypothesised structural relationships in the model.

8.6.3. Testing the Differences in Path Coefficients - the Australian Model versus the New Zealand Model

Significance testing for differences in path coefficients (standardised structural regression coefficients) for each hypothesised structural relationship ($= 0$ versus $\neq 0$) in the two models (Figure 8.2 and Figure 8.3) enabled the researcher to test her twelve hypotheses (Hc_1 through to Hc_{10c}) on the country effect. The test statistic t for difference for each hypothesised structural relationship was calculated (see equation 8.1) for significance testing under parametric assumptions, as prescribed in Rigdon, Ringle, and Sarstedt (2010).

$$t = \frac{k_{Aus} - k_{NZ}}{\left[\sqrt{\frac{(m-1)^2 * SE^2_{Aus}}{(m+n-2)} + \frac{(n-1)^2 * SE^2_{NZ}}{(m+n-2)}} \right] * \left[\sqrt{\frac{1}{m} + \frac{1}{n}} \right]} \quad (8.1)$$

Where *k values* (with country suffixes) represent the standardised structural regression coefficients, SEs (with country suffixes) represent the standard errors, m and n represent sample sizes for the Australian and New Zealand models respectively. The ‘degrees of freedom’ of each test statistic (the *t* value calculated via equation 8.1 for each structural relationship) is (m+ n -2).

Table 8.11 depicts the results on the twelve hypotheses on the differences in path coefficients. It is evident from the *p* values corresponding to the *t* values for difference in nine path relationships, that there is no evidence to reject the null hypothesis that ‘there is no difference in the path coefficients’, at 5% significance level (this is because, $p > 0.05$). As such, these nine hypotheses are supported by data, in that the independent (predictor) construct/s appears to have the same effect on the dependent construct (response) for Australia as for New Zealand. For three path relationships however, this was not the case (at 5% significance level)—the corresponding hypotheses being, *Hc*₂, *Hc*₄, and *Hc*_{10a}. The positive *t* values imply that path coefficients corresponding to these three hypotheses are *stronger for Australia*.⁴⁸

Table 8.11: The Results on the Hypotheses

Hypothesis	Path Relationships	Test (2-tail) for difference in path coefficients		Hypothesis supported?
		<i>t</i>	<i>p</i>	
<i>Hc</i> ₁	Mission → Strategy	0.441	0.660	Yes
<i>Hc</i> ₂	Mission → Financial Health	2.582	0.011	No
<i>Hc</i> ₃	Strategy → Financial Health	1.090	0.277	Yes
<i>Hc</i> ₄	Strategy → Org. Infrastructure	2.100	0.037	No
<i>Hc</i> ₅	Financial Health → Org. Infrastructure	1.577	0.116	Yes
<i>Hc</i> ₆	Strategy → People	1.217	0.225	Yes

⁴⁸ Strictly speaking, a one-tail *t* test for difference needs to be conducted.

Hypothesis	Path Relationships	Test (2-tail) for difference in path coefficients		Hypothesis supported?
		<i>t</i>	<i>p</i>	
Hc ₇	Financial Health → People	0.243	0.809	Yes
Hc ₈	Org. Infrastructure → Processes	1.093	0.276	Yes
Hc ₉	People → Processes	0.150	0.881	Yes
Hc _{10a}	Processes → Client Satisfaction	2.372	0.019	No
Hc _{10b}	Processes → People Satisfaction	0.581	0.562	Yes
Hc _{10c}	Processes → Donor Satisfaction	0.987	0.325	Yes

The unsupported hypotheses are restated as follows:

Hc₂: “Mission has an equal positive effect on the Financial Health in Australia as in New Zealand”.

Hc₄: “Strategy has an equal positive effect on the Organisational Infrastructure in Australia as in New Zealand”.

Hc_{10a}: “The Processes have an equal positive effect on Client Satisfaction in Australia as in New Zealand”.

The implications of the unsupported hypotheses, which as mentioned above translate to stronger path coefficients for Australia, are discussed in turn. However, the reader should note that the differences in path coefficients for the two countries are small (at 1% significance level the three hypotheses are supported, implying no difference), and therefore, may not have profoundly significant practical implications. For the convenience of the reader, the parameter estimates associated with the unsupported hypotheses are shown in Table 8.12.

Table 8.12: Parameters Associated With Unsupported Hypotheses

Path Associated with the Hypothesis <i>(the mediator causing an indirect effect, if present, is shown in square parenthesis)</i>	Path Coefficient (direct effect) <i>(the indirect effect through the mediator shown in parenthesis)</i>	
	Australia	New Zealand
Hc₂ : Mission → Financial Health [Strategy]	0.478 <i>(0.726*0.426 = 0.309)</i>	0.250 <i>(0.687*0.578 = 0.397)</i>
Hc₄ : Strategy → Organisational Infrastructure [Financial Health]	0.493 <i>(0.426*0.190 = 0.081)</i>	0.190 <i>(0.578*0.513 = 0.297)</i>
Hc_{10a} : Processes → Client Satisfaction [None]	0.802 <i>(not applicable)</i>	0.577 <i>(not applicable)</i>

Hc₂: Mission → Financial Health

As shown in Table 8.12, Mission has a stronger direct effect on the Financial Health for Australian healthcare NPOs (0.478) than for New Zealand healthcare NPOs (0.250). There is no such a significant difference in the indirect effect of Mission on the Financial Health, through the mediator ‘Strategy’ (0.309 for Australia and 0.397 for New Zealand). Therefore, it is clear that both the direct effect and the total effect (direct effect + indirect effect) of Mission on the Financial Health are higher for Australian NPOs than for New Zealand NPOs. The direct effect of Mission on the Financial Health, as mentioned earlier, signifies the Mission’s attractiveness to draw private and nongovernmental funding (Dolnicar, Irvine, & Lazarevski, 2008; Mount, 1996; Smith, Cronley, & Barr, 2012). Therefore, the researcher revisited the funding profile (Chapter Seven) of Australian healthcare NPOs and New Zealand healthcare NPOs separately. As observed in Chapter Seven (Figure 7.8), the level of government funding in both countries is similar. Also, the researcher was not able to attribute any other reason from her analysis of descriptive statistics (Chapter Seven), for a high direct and total effect for Australian healthcare NPOs. It could be possible that the observed increase is spurious. For example, the Australian managers could be over-confident that they have well-aligned Missions that can draw funding, when in fact there might not be any special mission-activeness in Australian NPOs over New Zealand NPOs. In any case, given the importance of the Financial Health to a NPO and the Mission’s potential to improve it, separate research must be undertaken to identify why the Mission seems to have a stronger direct and total effect on the Financial Health for Australian healthcare NPOs than for New Zealand healthcare NPOs.

Hc₄: Strategy → Organisational Infrastructure

As shown in Table 8.12, Strategy has a stronger direct effect on Organisational Infrastructure for Australian healthcare NPOs (0.493) than for New Zealand healthcare NPOs (0.190). However, the indirect effect of Strategy on Organisational Infrastructure (through the mediator Financial Health) is much less for Australian healthcare NPOs (0.081) than for New Zealand healthcare NPOs (0.297). Due to this compensation, the total effect of Strategy on Organisational Infrastructure is not substantially higher for Australian healthcare NPOs (0.574, being 0.493 + 0.081) than for New Zealand healthcare NPOs (0.487, being 0.190 + 0.297). However, the difference seems to be nontrivial still, suggesting that the Strategy to Organisational Infrastructure relationship is stronger for Australian healthcare NPOs than for New Zealand healthcare NPOs. One possibility for this discrepancy could be the relatively larger size of Australian healthcare NPOs, as shown in Chapter Seven (see Figure 7.9). A future study may be undertaken to examine any discrepancies more closely.

Hc_{10a}: Processes → Client Satisfaction

The relationship between Processes and Client Satisfaction is not mediated by a third variable. The path coefficients depicted in Table 8.12 clearly show that Processes are more strongly related to Client Satisfaction in Australian healthcare NPOs than in New Zealand healthcare NPOs. This is also evident from the R^2 values associated with Client Satisfaction, being 0.643 (square of the path coefficient 0.802) and 0.344 (square of the path coefficient 0.577) for Australia and New Zealand respectively. This is indeed a marked difference. In an ideal world, and all other things being equal, this difference should be attributable to the greater efficiency and effectiveness of the Processes of Australian healthcare NPOs, relative to New Zealand healthcare NPOs. If this was the main reason for the difference, then it pays for senior managers in New Zealand to learn from their Australian counterparts why and how the Australian healthcare NPOs are performing better in satisfying the clients. It is important to note that process efficiency and effectiveness could be a function of economies of scale, with larger NPOs, mostly found in Australia (see Figure 7.9 in Chapter Seven), being able to utilise their resources better in the production of client services (Folland, Goodman, & Stano, 2007; McWilliams & Siegel, 2001).

The above said, it could be possible that Australian managers, as mentioned earlier, are somewhat coloured by their convictions, in that they could be overconfident that they have done things right (Trevor-Roberts et al., 2003). In designing future research to identify why Australian healthcare NPOs are seemingly better at satisfying the clients (a stronger Processes → Client Satisfaction relationship), care should be taken to eliminate any spurious effect such as the size of the NPO (economies of scale). Obtaining independent data on Client Satisfaction could be considered in a future study.

8.7. RESULTS AND DISCUSSION OF THE EMPIRICAL WEIGHTS FOR THE PERFORMANCE DIMENSIONS

In this section, as mentioned in the chapter introduction, the researcher sets out and discusses her results on the ancillary model (the development of this model was covered in section 4.8, Chapter Four) to achieve her fourth research objective of determining empirical weights for each performance dimension (and its subparts) to enable NPOs to assess their overall (strategic) performance. As shown in Chapter Four (see Figure 4.7 for example), the Overall Performance Index (OPI) was constructed using 6 dimensions and 13 sub-dimensions: 1. Mission (1.1 Social Profit, 1.2 Organisational and People Excellence); 2. Strategy (2.1 Strategy Planning, 2.2 Strategy Implementation); 3. Capabilities and People Development (3.1 Organisational Infrastructure, 3.2 People); 4. Financial Health (4.1 Cost Reduction Initiatives, 4.2 Revenue Increasing Initiatives); 5. Processes (5.1 Knowledge Management and Continuous Improvement, 5.2 Process Design); and 6. Stakeholder Satisfaction (6.1 Client Satisfaction, 6.2 People Satisfaction, 6.3 Donor Satisfaction).

The results in this section are limited to parameter estimates of the ancillary model. Like in any other structural equation model, the quality of the ancillary model needs to be assessed before model parameters could be interpreted. Appendix E shows the results on ‘scale reliability’ and ‘construct validity’, using the exact same criteria that were used to assess the quality of the researcher’s main model. Results in Appendix E indicate that the quality of the ancillary model is satisfactory for the purpose of interpreting parameter estimates.

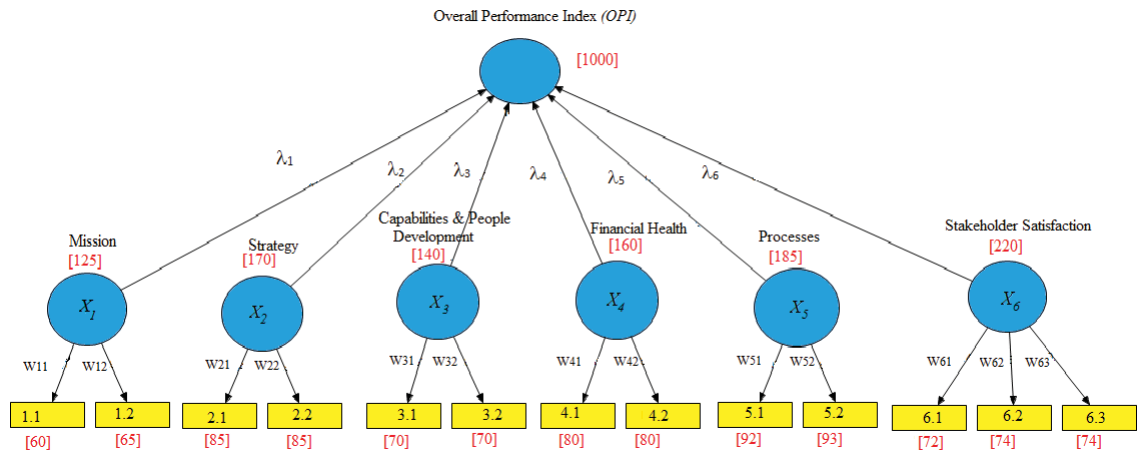


Figure 8.4: The hierarchical model on overall performance

As mentioned in section 4.8, testing the ancillary model and expressing the OPI as a weighted linear combination of the constructs was accomplished by analysing the regression coefficients (λ_i) estimated. Table 8.13 depicts the standardised regression coefficients of the constructs and their subparts/items (also see Figure 8.4).

In establishing the scoring system, as mentioned earlier, 1000 points were assigned to overall performance, to be consistent with scoring systems used in PM systems such as the Baldrige Excellence Framework and the EFQM Excellence Model. The, 1000 points were apportioned across the six constructs (Figure 8.4) linearly, based on the standardised regression coefficients (λ_i) estimated for respective constructs. Thus the formula that was used to assign points to each construct is shown in equation 8.2 below.

$$\text{Score for construct (measurement category)} \quad i = 1000 * \lambda_i / \sum_{i=1}^6 \lambda_i \quad (8.2)$$

Once the score for each measurement category was determined, the category score was apportioned across the respective measurement items (subparts) using the item weights (w_{ij}) estimated through the PLSBSEM procedure (Table 8.13), using a procedure similar to what was used to apportion 1000 points across the six measurement categories. Table 8.14 depicts the estimated scores for each category and measurement item, and therefore, the scoring system to assess overall organisational performance. For the convenience of the reader, the quantitative survey questionnaire items (survey

questions) that operationalise the constructs and their subparts are also shown in Table 8.14.

Table 8.13: Standardised Regression Coefficients of Constructs and Items

Constructs (Measurement Categories)	Standardised Regression Coefficients (λ_i)	Weights of the Measurement Items (w_{ij})		
Mission ($i = 1$)	0.140 (λ_1)	0.783	0.902	
Strategy ($i = 2$)	0.190 (λ_2)	0.94	0.929	
Org. Capabilities and People Development ($i = 3$)	0.158 (λ_3)	0.884	0.853	
Financial Health ($i = 4$)	0.180 (λ_4)	0.894	0.913	
Processes ($i = 5$)	0.208 (λ_5)	0.959	0.961	
Stakeholder Satisfaction ($i = 6$)	0.247 (λ_6)	0.844	0.852	0.853

Table 8.14: The Summary Description of the Point Scoring System

Measurement Categories	Measurement Items (Subparts of the Construct)	Relevant Survey Questions	Estimated Points	
			Item	Category
1. Mission	1.1 Envisioning Social Profit	Q1, Q5	60	125
	1.2 Envisioning Organisational and People Excellence	Q9, Q13	65	
2. Strategy	2.1 Strategic Planning	Q16, Q17	85	170
	2.2 Strategy Implementation	Q24, Q18	85	
3. Organisational Capabilities and People Development	3.1 Organisational Infrastructure	Q34, Q20	70	140
	3.2 People	Q29, Q19	70	
4. Financial Health	4.1 Cost Control	Q2, Q15	80	160
	4.2 Cash Flow Increase	Q3, Q12, Q14	80	
5. Processes	5.1 Continuous Improvement and Knowledge Management	Q23, Q36, Q40	92	185
	5.2 Well Designed, Safe, Efficient and Effective Processes	Q25, Q35, Q37	93	

Measurement Categories	Measurement Items (Subparts of the Construct)	Relevant Survey Questions	Estimated Points	
			Item	Category
6. Stakeholder Satisfaction	6.1 Client Satisfaction	Q6, Q8, Q26, Q32	72	220
	6.2 People Satisfaction	Q7, Q21, Q28, Q33, Q41	74	
	6.3 Donor Satisfaction	Q11, Q22, Q27	74	
Overall Performance			1000	1000

8.7.1. Discussion of Empirical Results on the Ancillary Model

Kaplan (2008) admitted that the BSC lacks an overall performance measure; in relation to the for-profit BSC, Kaplan opined that an overall performance measure should be an unbiased/balanced measure that can be used by shareholders for “executive compensation”. Similar reasoning can be given for an overall performance measure for NPOs, for the purpose of funding, governance, and accountability. This is because major donors would be more concerned about the overall performance of their recipient organisation, because the overall performance is a better indicator of the long-term performance of a NPO than a single measurement category such as Financial Health (Herman & Renz, 2004; Ritchie & Kolodinsky, 2003). One of the most important practitioner related outcomes associated with developing and testing the ancillary model is the introduction of the *scoring system for NPOs*. Since the performance measures of a nonprofit BSC are unweighted, the BSC fails to explain how the measures, as a whole, indicate how the organisation is progressing overall, in implementing its strategies. For a NPO, this means that key donors (especially the government) are unable to gauge easily how the leadership (the board of directors) is performing, to link funding to performance. In keeping with “statistical thinking” (Hoerl & Snee, 2012), consolidating (aggregating) each performance measure in a PM system *on a rational basis*, to constitute an aggregate measure on performance (e.g. performance in the Strategy, Overall Performance) will filter measurement error and common cause variation (noise). As an example, a score related to a particular survey question on Strategy (say Q16), is more prone to noise than an aggregated score for Strategy, which represents five survey questions (see Appendix C). This is not to undermine the fact that managers need to be able to distinguish random variation (as indicated by an individual performance measure) from special cause variation, through statistical literacy.

Statistical process control tools such as run charts and control charts should not be overlooked in performance monitoring.

The scores in Table 8.14 suggest that some constructs are more important than others. For example, the final construct *Stakeholder Satisfaction* carries a weight (regression coefficient) of 0.247, but Mission, which is the driver of performance (in the context of the researcher's main model), carries a weight of only 0.140. Thus, when empirically distributing the total score (weight) of 1000 points allotted to the OPI, *Stakeholder Satisfaction* gets 220 points but Mission gets only 125 points (empirical weights of all PM dimensions and their subparts are given in Table 8.14). From a practical perspective, the differences in empirical weights (hence the scores) estimated for the constructs can be attributed to the breadth and depth covered in the constructs. For example, it becomes clear from Table 8.14 that the number of survey questions assigned to each construct differs considerably. Since Mission has a narrower domain (just four survey items were required to cover Mission) than Stakeholder Satisfaction (as many as twelve survey items were required to cover Stakeholder Satisfaction), it is reasonable to believe that Mission would command less weight than Stakeholder Satisfaction, which covers the satisfaction of three diverse groups of stakeholders. In addition, Stakeholder Satisfaction itself covers achievement of the Mission.

Very little is known about causal factors that contribute to the success of NPOs. Besides, there is substantial heterogeneity between NPOs. Under these circumstances, use of alternative methods such as multi-criteria decision making techniques have limited scope in developing a performance scoring system for NPOs. The researcher believes that her approach of fitting a performance index model to empirical data provides a useful and unbiased performance scoring system that NPOs can use as a dashboard to monitor and improve their performance. One can also view the researcher's findings as further evidence of the robustness of the BSC.

The scoring system that the researcher developed is useful to NPOs for self-assessment and performance improvement, because a scoring system keeps an organisation in check. As described earlier, the researcher's quantitative questionnaire can be used as the starting point of self-assessing the score for each measurement item, category, and overall performance (see Table 8.14). The reader should note that what is shown in

Table 8.14 is an extremely condensed version of the questionnaire contents (the full questionnaire is shown in Appendix C). One of the most important requirements in self-assessment is to make a reliable assessment of each item being measured. The reliability of assessment can be increased if the assessment can be made by the management team, first individually, and then collectively, for moderation. It is well known that self-assessment encourages continuous improvement, as it allows teamwork in solving high-level (strategic) problems (Samuelsson & Nilsson, 2002). The researcher's scoring system is also useful for large donors such as government agencies that provide substantial funding to NPOs. Large donors typically provide funding for multiple beneficiaries, and the scoring system can be used to monitor the relative performance (i.e. comparative benchmarking) of their beneficiaries. Moreover, a scoring system will also force NPOs to become more accountable to their key stakeholders.

The researcher believes that the scoring system that she developed and the accompanying self-assessment questionnaire would help NPOs, because her work is based on empirical research, rather than universal prescriptions. As such, the basis of the findings is the researcher's theory on nonprofit performance and confirmation by the data, collected from a large number of NPOs in Australasia (albeit in healthcare). Within a PLSBSEM framework, the researcher's measurement model fitted the data well—the AVE values of the measurement categories were high (see Appendix E). Besides, in PLSBSEM, the variability of the second order construct (in the researcher's case, the OPI) is fully explained by its first order constructs (Chin, 1998; Hair et al., 2014). This basically means that the researcher's results will be generalisable, although there will be exceptions (outlier organisations).

Having determined the empirical weights for each performance dimension (and its subparts) to enable NPOs to assess their overall strategic performance, the researcher has demonstrated that she has achieved her fourth research objective.

8.8. IMPLICATIONS FOR MANAGEMENT PRACTICE AND PERFORMANCE MEASUREMENT

The researcher's PM system (the main model, the ancillary model and the operationalisation of the constructs) helps managers of NPOs to align Mission, Strategy, Strategic Resources (People and the Organisational Infrastructure) and Processes to achieve the Missions of their NPOs, by being able to satisfy the key Stakeholders (Clients, People and the Donors). This is because the PM system predicts and explains how Mission and Strategy drives the system (Strategic Resources and the Processes) to deliver Mission-related outputs and outcomes (researcher's main model). The researcher defined the constructs of her main model nominally as well as operationally. The researcher's main model contains nine constructs: Mission, Strategy (the two drivers that drive the strategic resources); Financial Health, People, Organisational Infrastructure (the three Strategic Resources that drive the Processes); Processes; Client Satisfaction, People Satisfaction, Donor Satisfaction (the three multifaceted outcomes of the Processes). The researcher's models (the main model and the ancillary model) have implications for management practice on two fronts. The first is on the operationalisation of the constructs (PM dimensions) and the second is on predicting and explaining how performance can be improved. These are discussed below.

8.8.1. Operationalisation of the Constructs

Operationalisation of the two drivers Mission and Strategy are very important for NPOs. In articulating the theoretical basis of the nonprofit BSC, Kaplan (2001) asserted that the Mission offers a sense of purpose of the NPO, and therefore, the Mission should be the starting point in PM. This implies that if the Mission is linked poorly to the Strategy, and the Strategy is not well deployed in the organisation, the outcomes (satisfaction of the key Stakeholders and thereby, achieving the Mission) are not going to be outstanding. However, the nonprofit BSC does not operationalise Mission and Strategy. The researcher's model, however, does. Since the researcher's model operationalises Mission (questionnaire items Q1, Q5, Q9 and Q13), senior managers in NPOs can measure how well the Mission is being approached and deployed in their organisations. In particular, managers can examine to what extent the Mission achieves the following: (a) strives for effective social services (Q1), (b) strives to create independence, dignity and opportunities for clients to improve their living (Q5), (c)

creates service leadership (Q9), and (d) creates people development to service the clients (Q13). The role of senior management in a NPO, as far as PM is concerned, is to design appropriate performance measures that capture these requirements, given the Mission.

Similar to Mission, since the researcher's model does operationalise Strategy (questionnaire items Q16, Q17, Q18, Q24 and Q39), senior managers in NPOs can measure how well the Strategy is being approached and deployed in their organisations, through appropriate performance measures. The operational definition of Strategy emphasises the following: incorporating stakeholder needs in strategy formulation (Q16), leadership engagement (Q17), partnering and collaboration in social responsibility (Q24), strategic resource identification and creating the organisational structure to implement strategies (Q18), and matching operating plans to employees to motivate the employees (Q39). Again, the role of senior management in a NPO, as far as PM is concerned, is to design appropriate performance measures that capture these requirements. While the need to have strategies to motivate people was something that was identified in the fieldwork (the case study), this is well supported in the literature. Over four decades ago, in their seminal work on job enrichment, Hackman and Oldham (1975) articulated the concept "experienced meaningfulness of work": designing work in such a way that employees know how their work contributes towards achieving organisational goals. The researcher observes that this is still one of the key factors that motivates people to perform (e.g. see May, Gilson, & Harter, 2004; Peng et al., 2016; Woods & Sofat, 2013).

The researcher's three constructs that deal with strategic resources, namely Financial Health (operationalised via Q2, Q3, Q12, Q14, and Q15), People (operationalised via Q19, Q29, Q30), and Organisational Infrastructure (operationalised via Q20 and Q34) help senior managers to devise performance measures to examine how strategic resources are being used in their organisation. One could contend that the strategic resource constructs in the researcher's main model are actually embedded in the BSC (nonprofit version or otherwise), and therefore, there is nothing new in these resource constructs. For example, People and Organisational Infrastructure constructs can be amalgamated as *Learning and Growth*, for practical purposes. Although amalgamating constructs to reduce the number of PM dimensions, makes the model simple, this would

come at the expense of losing the explanatory power. The BSC needs the strategy map to provide this explanatory power. Managers in NPOs can avoid the difficult task of drawing strategy maps to identify causality between strategic objectives, by using the researcher's main model (Bisbe & Barrubés, 2012). Drawing strategy maps does not directly help managers in determining performance measures, because according to the BSC, performance measures eventually come from strategic objectives (the performance measure makes the strategic objective measurable). Kaplan's assertions that (in a nonprofit context) Strategy and Mission should feature in the BSC and that Mission should drive the Strategy are embedded in an organisation's strategy map, but not in the implied theoretical relationships between the BSC's PM dimensions. In the researcher's main model, the operational definitions of constructs (all the constructs including the resource constructs) help managers in designing the performance measures, in conjunction with the strategic objectives; as explained earlier (see section 6.7.2 in Chapter Six), the strategic objectives play a role in translating the critical success factors (CSFs) into key performance indicators (KPIs). For example, coming back to Q2, the first aspect of Financial Health, the senior managers of a NPO can ask for themselves: *“what would be a good KPI that reflects our investments in IT to create a paperless culture to improve productivity?”*

Another important ‘operationalisation related implication’ of the researcher's main model to management practice is that her model uses separate constructs to directly capture satisfaction of the key stakeholders: Clients, People, and Donors. This makes Stakeholder Satisfaction measurement mandatory, and reminds managers that eventually, the signal (feedback) to improve Processes (also to better utilise strategic resources) originates from the stakeholders (more precisely the level of satisfaction of the stakeholders).

8.8.2. Predicting and Explaining how Performance can be Improved

The prediction and explanation aspect embodied in the researcher's main model, and its implication for management practice, was partially explained earlier (section 8.4). That is, because the model explains (through causal linkages) how a NPO achieves Stakeholder Satisfaction, and thereby the Mission, it can be used as a tool to explain successful performance improvement as well as unsuccessful performance

improvement. The model can be used to provide a clear picture about how their strategic objectives are conveyed and communicated throughout the organisation, without having to use a strategy map. In a BSC framework, the strategy map accomplishes this goal but the issue is that preparing a strategy map becomes a monumental task (unless external consultants are invited to perform this task) for many NPOs, especially in the case of identifying causal connections between strategic objectives (Atkinson, 2006; Banker, Chang, & Pizzini, 2011; Bisbe & Barrubés, 2012). In the researcher's model, causality comes from her theory, and the theory lies outside the scope of causal maps.

The second benefit of the strong causal emphasis in the researcher's main model is that it facilitates feedback. Needless to say, if A causes B, then if the performance of B is low, then this sends feedback to managers that A should be improved to improve B. Therefore, in the researcher's main model (also see her conceptual model), feedback flows from right to left, starting from Stakeholder Satisfaction. In essence, the feedback mechanism provides information on whether or not the organisation achieves its Mission, by satisfying its key stakeholders. Thus, managers can make the necessary interventions on a regular basis, by improving the Processes (the causes of Stakeholder Satisfaction). To improve the Processes, managers need to improve the Financial Health, People, and the Organisational Infrastructure. To improve these strategic resources, the organisation (system) has to be driven by its Mission, via an effective Strategy. This is essentially what the causal links in the researcher's main model and the implied feedback mechanisms indicate.

The model also reminds managers that there are three key stakeholders in NPOs and that each stakeholder has their own expectations (this is why three separate constructs had to be used to measure Stakeholder Satisfaction). An organisation needs to design its Processes based on what works best for each stakeholder (via stakeholder feedback), rather than being too preoccupied by serving the community or the donors alone. The main model emphasises the importance of employee skills (via the construct People) and the information system (via the construct Organisational Infrastructure) to deliver the stakeholder outcomes. For example, in understanding the clients' expectations about quality, an effective service organisation has to excel in its Processes, which in turn requires employee training and an effective information system to enhance

organisational learning (Argote, 2012; Levitt & March, 1988). The researcher's main model also shows the importance of the Financial Health of the NPO to execute its Strategy, and the role Mission plays in improving Financial Health.

8.8.3. Implication of the Ancillary Model

One of the most important benefits of the researcher's ancillary model is that it provides a scoring system for managers to quickly evaluate how the organisation is performing in each performance dimension (construct or category), as well as overall, without having to use performance measures (KPIs). As argued by the researcher earlier, aggregation of performance indicated by each performance measure is problematic. The researcher's survey questionnaire (Appendix C) can be used to collect the quantitative data, and the scoring system can be used to self-assess the performance of a NPO, in order to identify areas for improvement. One of the most important aspects in a self-assessment is to make an accurate assessment of each PM item being measured. The researcher believes that her approach of fitting a performance index model to empirical data provides a useful and unbiased performance scoring system that NPOs can use as a dashboard to monitor and improve their performance. From a practical perspective, by making an accurate assessment of each PM item, an organisation can get an idea about their strengths and weaknesses. Moreover, donors such as the government agencies that provide substantial funding to NPOs can use the scoring system to monitor the relative performance of their beneficiaries. It is also valuable for organisations to plan their future improvements and related strategies. A NPO can use the following 2-step procedure to obtain a representative score for each PM item (sub-dimension), category (dimension), and the overall performance.

Step 1: The chairperson (or their representative) of the NPO requests the senior managers to individually respond to the questionnaire (no collusion allowed in step 1) using the questionnaire provided (Appendix C); the managers need to be advised that the score of a measurement item (the subpart or sub-dimension of the measurement category) will be calculated by taking the arithmetic mean of the questionnaire responses pertaining to each item. For example, if a manager assigns a 6 (Agree) to Q1 and a 5 (Agree) to Q5 in the Likert scale being used, the score given by the manager for item 1.1 "Envisioning Social

Profit” should be taken as 5.5 (see Table 8.14). It is important that the chairperson should also respond to the questionnaire.

Step 2: The chairperson (or their representative) collects the responses and determines the consensus score for each item, through a consultation process (a meeting could be called to agree upon on the consensus score, as it is highly unlikely that all managers would give the same rating to the same questionnaire item).

Once the consensus scores for all 13 items (sub-dimensions) have been determined, these scores can be converted to points by linear scale transformation, using the following logic: a consensus score of 1 (Strongly Disagree) equates 0 points and a consensus score of 7 equates to the (maximum) points assigned to the item (see Table 8.14). Therefore, if the consensus score for an item is N (in the 1 to 7 Likert scale), and the (maximum) points assigned to that item is M , the point score X for that item is given by the following equation (Eq 8.3).

$$X = \frac{M}{(7-1)} * (N - 1) \quad (8.3)$$

To demonstrate the use of equation 8.3, say if the consensus score for item 1.1 is 5.5 (in the 1 to 7 Likert scale), then the points score for item 1.1 would be 45 points, being $(60/6) * (5.5 - 1)$. Once the score for each item has been computed, managers can calculate the scores for the measurement categories and the Overall Performance. The Overall Performance score would be the sum of all 13 item scores.

The next step would be to place the Overall Performance score in perspective (i.e. grade the score), much the same way as an academic performance is graded using a rational criterion. As similar to the guidelines given by the advocates of excellence models (e.g. Brown, 2013; EFQM, 2013), overall scores in the proximity of 100%, 75%, 50%, 25%, and 0% can be interpreted as Outstanding, Very Good, Good, Poor, and Unfathomable respectively. An overall score of 100% (1000 points) is impossible to achieve, and so is achieving 0% (zero points); a no score implies that the organisation being evaluated has no performance management system (Brown, 2013). According to Brown, 500 points (50%) is considered as above average performance (though 50% is the nominal middle

point), meaning that for a large number of organisations, scoring beyond 500 points is a difficult proposition (i.e. more technically, median and mode < 500 points). Brown also contends that an organisation that scores between 75% and 100% are often sufficient to win “quality awards” at national level. Brown also asserts that an organisation that scores below 250 points (25%) should consider radical improvements, involving substantial change management interventions for better outcomes. The important feature of the self-assessment process is that without any special training, managers can evaluate the strategic performance of their organisation. It is important to note that the score is just a number, indicating a performance gap; how this gap should be bridged is a managerial decision; the researcher’s other contributions—the main model, operational descriptions of the PM dimensions, and the CSFs in particular—could help managers in bridging performance gaps.

The researcher cautions that an organisation should not spontaneously react to a low item score in the survey, as long as it has scored well in the other items belonging to the same category (see Table 8.14). This is because a low score can sometimes be a consequence of incomplete information. For example, an organisation should not be overly concerned if they score low in Q6, as long as it scores well in the remaining questions under Stakeholder Satisfaction. Therefore, in general, category scores will be more robust than item scores, which in turn will be more robust than a single survey questionnaire item score.

It is important to note that the scoring system derived from the researcher’s ancillary model for self-assessment should always be used in conjunction with the researcher’s main model. This is because it is the main model that predicts and explains how strategic objectives are achieved. In other words, the main model shows the causal links between the PM dimensions. If the performance in one PM dimension is low, by applying a causal-predictive logic, managers can identify what PM dimensions are causing this low performance. For example, if a NPO is not performing well in the Financial Health dimension, managers in that organisation will know that this has to do with the Mission and Strategy, because it is these two PM dimensions that cause the Financial Health (Figure 8.1). Managers should use appropriate performance measures for each construct, using the operational descriptions and CSFs provided by the researcher (see section 6.7.2 in Chapter Six), in conjunction with the strategic objectives

of their organisation. The researcher does not insist that a NPO should have a performance indicator to match each survey questionnaire item. The questionnaire contained 41 survey items, of which Q4 should be excluded because it is the global measure of Financial Health. In addition, Q10, Q30, Q31, Q38, and Q39 should also be excluded due to reliability and validity reasons mentioned elsewhere. The remaining 35 questionnaire items become candidates for generating performance measures. Based on the strategic objectives of the NPO, managers can combine some questionnaire items on a rational basis to reduce the number of performance measures for their organisation.⁴⁹

In summary, the self-assessment method prescribed by the researcher (the scoring system) and the main model will help a NPO to: (a) identify (or at least look for) best practices for improvement, (b) identify the current trends in performance management practices, (c) engage in benchmarking for planning, implementing and reviewing its strategies, and (d) establish better stakeholder relationships to achieve the Mission. With this, the researcher concludes her discussion on explaining how she achieved the general research objective (i.e. developing guidelines for practitioners in NPOs to help them assess the strategic performance of their organisations).

8.9. COMPARING THE RESEARCHER'S PM MODEL WITH THE NONPROFIT BSC AND ORGANISATIONAL EXCELLENCE MODELS

Kaplan (2001) argues that a NPO should include its Mission in designing their scorecard, since Mission reflects the long-term objectives of the NPO. The researcher's main model is consistent with Kaplan's assertion. Both the nonprofit BSC and the researcher's model give top priority to organisational Mission since NPOs are mission driven organisations; that is, Mission is the reason for the existence of NPOs and not maximising the wealth (net benefits of future returns) of the shareholders, as in the case of profit making organisations. In the nonprofit BSC, the Mission is posited to drive the Strategy rather than the shareholder objectives (Kaplan, 2001), but neither the Mission nor the Strategy is measured explicitly. In the researcher's model, both Mission and Strategy are explicitly measured. Essentially, the researcher's main model is more explicit on the nonprofit BSC's notion that the Mission drives the organisational Strategy, which in turn drives the system inputs: all the tangible and intangible

⁴⁹ The BSC envisages about 3 to 5 measures in each PM dimension, thus about 15 to 25 performance measures in total (Lawrie & Cobbold, 2004).

resources. The BSC labels these resources and Learning and Growth but the researcher uses three PM dimensions to reflect an organisation's *human capital* (the People dimension), *organisational capital* (the Organisational Infrastructure dimension), and the *financial capital* (the Financial Health dimension).

One of the prominent features of the current version of the BSC is the *strategy map*. The strategy map shows how each strategic objective is causally related to one another to achieve, in the case of a NPO, the Mission. Placing each strategic objective in the appropriate BSC dimension and generating performance measures for each strategic objective occurs as the strategy map is being developed. Kaplan (2008) argued that by including the strategy map, the BSC becomes more directly involved in strategic planning. Stated alternatively, Kaplan contends that the BSC becomes an important consideration in strategic planning. While the researcher does not dispute this—and also acknowledges that a strategy map promotes systems thinking—the literature review (e.g. Banker et al., 2011; Bisbe & Barrubés, 2012) as well as her fieldwork (the case study) found that designing strategic maps is very challenging and many NPOs do not use them. A NPO could invite external consultants to help them in preparing a strategy map and other requirements of the BSC, but there is no evidence to show that NPOs actually use external consultants for this type of activity. The organisational culture and financial pressures may prohibit most NPOs from hiring external consultants for strategic-level activity. The researcher's model avoids the use of a strategy map. As argued by the researcher earlier, strategy maps do not directly help managers in designing performance measures because performance measures come from the strategic objectives, which in turn come from strategic planning. If the practitioners do not use strategy maps in strategy planning, they become less useful. This is why a generalisable theory linking PM dimensions (along with the operational definitions of PM dimensions) was developed by the researcher (her main model) in the first place.

The literature shows that measuring success of NPOs is challenging, because success should be captured through the satisfaction of stakeholders, including donors, staff, volunteers, and clients who may have different expectations (Balser & McClusky, 2005; Kendall & Knapp, 2000; MacIndoe & Barman, 2013). The researcher's study confirmed this. To accommodate the BSC framework for NPOs, Kaplan (2001) expanded the definition of the customer (for the nonprofit BSC) to cover both donors as

well as the clients of the organisation; achieving customer results (customer satisfaction) was seen as achieving the Mission. The researcher's PM model looks at the Mission achievement through satisfaction of 3 key stakeholders: clients, employees, and donors (the researcher developed measures for all three of them).

The researcher's PM model shows the links between Mission, Strategy and the remaining elements of the system (inputs, processes, and outputs). While the BSC emphasises the importance of aligning organisational Mission with Strategy, which in turn feeds (via the strategy maps) the remaining elements of the system embedded in the PM dimensions, the BSC does not provide a clear causal connection between these elements. Therefore, the researcher's model can be considered as an extension of the BSC for NPOs. The researcher's PM model also plays the role of an organisational performance theory, because it explains how the organisation adds value to the community or the targeted group, since it measures Stakeholder Satisfaction (hence mission achievement) by aligning the organisational Mission, Strategy, Resources (Financial Health, People, and the Organisational Infrastructure) and Processes to one another.

The BSC's Internal Process perspective is operationalised through two sub-dimensions: Critical Processes and Innovation Processes (Kaplan, 2001). Apart from these sub-dimensions, the researcher's PM model captures the third sub-dimension Continuous Improvement, which reflects the incremental approach in the quality literature.

The BSC's Learning and Growth perspective arises from two sub-dimensions: People and Systems (Kaplan, 2001; Kaplan & Norton, 2004). The parallel concept embodied in the researcher's main model 'Organisational Capability, Infrastructure, and People Development' (in the opinion of the researcher, this is a better representation of BSC's Learning and Growth perspective) is operationalised through two dimensions: People and Organisational Infrastructure. All of the features of the researcher's model mentioned provide further evidence that the contributions of the researcher's study come from the extension of the nonprofit BSC.

The EFQM Excellence Model and the Baldrige Excellence Framework are the dominant organisational excellence models (Al-Tabbaa et al., 2013; Grigg & Mann, 2008) used

by organisations (NPOs included). Both excellence models emphasise the importance of People and consider people as an Enabler of Results through process improvement (Neely, 2007; NIST, 2015; Paton, 2003; Srikanthan & Dalrymple, 2003); thus these models highlight the importance of employee contribution for organisational success and sustainability (Al-Tabbaa et al., 2013; Neely, 2007). Similar to excellence models, the researcher's PM model also considers the people to be a major strength of the organisation and highlights the importance of their development and satisfaction. One difficulty that managers face in adopting excellence models in NPOs is the Customer criterion. This criterion should be modified to 'Stakeholders', to be more applicable to the sector (Al-Tabbaa et al., 2013).

A more apparent difference between the researcher's PM model and the excellence models is in the Driver. The researcher's PM model posits the Mission as the driver of performance. As mentioned earlier, the study was guided by the conceptual model, for which the nonprofit BSC (Kaplan, 2001) provided a significant theoretical basis. According to Kaplan, the strategy of a NPO is driven by the Mission, and without a Mission, there is no such thing as a NPO. In contrast, the strategy of a profit making organisation (e.g. growth, market share), according to Kaplan and those who embrace the shareholder view, is driven by shareholder interests, which, the leadership is supposed to safeguard—as their agents—implying that Leadership is the driver of a for-profit organisation. The excellence models (originally developed for for-profit organisations, but, later extended to NPOs) take the position that Leadership drives the system to achieve the results (Collier & Evans, 2015), which is consistent with the for-profit BSC. Closer examination of the operational descriptions of the Mission implies that the researcher is also representing some aspects of Leadership under the label Mission. For example, consider the fourth operational description of Mission and Core Values: “...pursue leadership in what an NPO does to energise staff”. This is what could be considered as “transformational leadership”, which is essential for contemporary organisations (Bass, 1990; Haley, 2014). In his exhaustive review of PM systems, Evans (2004) argued that the difference between the BSC and excellence models is only in the semantics. Thus, proponents of excellence model can argue that whether Leadership or Mission should be the driver of a NPO could be in the semantics. Yet another difference between the researcher's PM model and excellence models is in the operational definitions and the points scoring system. Excellence models do not

distinguish between for-profit and NPOs. They use the same constructs (PM dimensions), same framework (theory), and same scoring system across for-profit and NPOs. The researcher questioned the validity of this stance. The PM system developed by the researcher is meant for NPOs only (however, the scoring system and the operational descriptions of the constructs were derived from data collected from the healthcare sector).

By comparing researcher's model with the nonprofit BSC and excellence models, the researcher has achieved her fifth research objective.

8.10. CHAPTER CONCLUSION

Arguably, this chapter is one of the most important chapters of the thesis, because it is in this chapter that all research questions have been comprehensively answered (RQ1 through to RQ3 were partially answered in Chapter Six). In this chapter, the researcher presented the test results of her main theoretical model and the ancillary model and discussed these results from a theoretical and practical standpoint. In testing her main theoretical model empirically, the researcher was able to achieve the following.

- ✓ Confirm the validity of the PM dimensions (constructs) of her model.
- ✓ Confirm the validity of the operational definitions of the PM dimensions.
- ✓ Show that the twelve hypothesised relationships (these predict and explain strategic performance of a NPO) were supported by the data.

Confirming the validity of the PM dimensions (section 8.2) demonstrated that the PM dimensions in the integrated PM framework developed by the researcher, are valid. This answered RQ1. Confirming the validity of the operational definitions of the PM dimensions (developed through the qualitative study) demonstrated that these operational definitions were supported by data.⁵⁰ This answered RQ2. Showing that the twelve hypothesised relationships were supported by the data basically equates to showing that the researcher's hypotheses stay in contention. This in effect answered RQ3. In addition, interpretation of the results on the hypothesised paths (more

⁵⁰ The exceptions are Q10, Q30, Q31, Q38 and Q39. These items should not be included in the operationalisation of the constructs.

technically the path coefficients) also provided further evidence of the validity of the researcher's main model.

Since RQ1, RQ2, and RQ3 are tied to the first research objective (Figure 8.5), and the fact that the researcher answered these research questions comprehensively, means that the researcher was able to achieve the first research objective successfully. Likewise, since RQ3 is directly tied to the third research objective and the fact that the researcher was able to answer this research question comprehensively, means that the researcher was able to achieve the third research objective successfully.

Confirming the validity of the ancillary model (section 8.7) paved the way to estimate the weights managers need to assign to each PM dimension and its subparts, in assessing the overall strategic performance of their NPOs. This answered RQ4 and in effect, the researcher was able to achieve the fourth research objective (Figure 8.5). In discussing the practical aspects of the weights, the researcher explained how her scoring system can be used by managers to improve the performance of their NPOs through self-assessment, to identify strengths and weaknesses. Based on the scoring system and the researcher's main model, the researcher proposed guidelines for nonprofit practitioners to assess the strategic performance of their organisations. Through this, the researcher was able to achieve her general research objective. Finally, the researcher achieved the fifth research objective based on the findings associated with research objectives 1, 3 and 4 in conjunction with the literature (details in section 8.9; also see Figure 8.5).

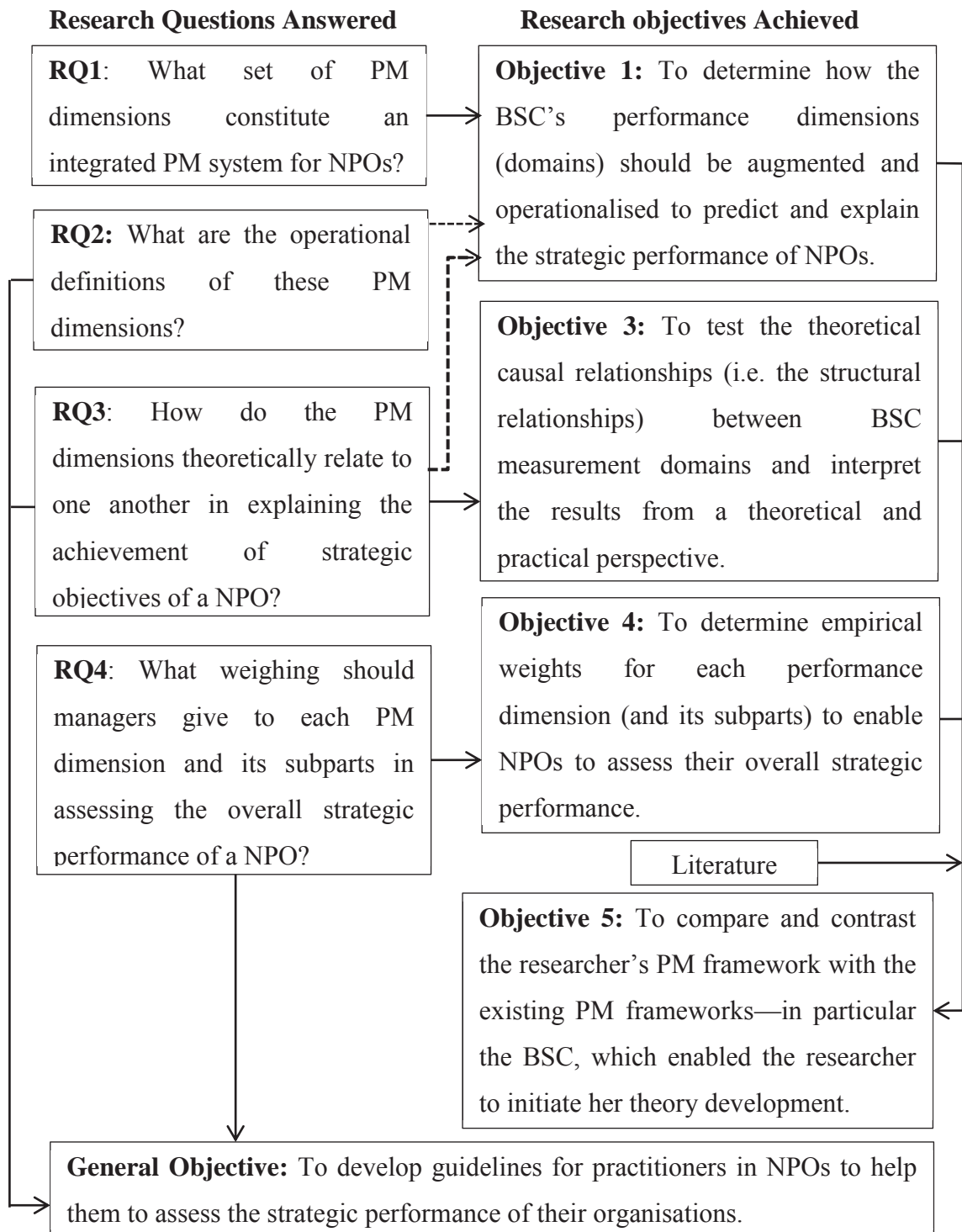


Figure 8.5: Research questions answered and research objectives achieved

The next chapter (Chapter Nine), which is the final chapter of the thesis, provides a summary of the research, explaining basis of the aims and objectives of the study and how each objective was achieved. The *new knowledge* contribution and further research directions are also covered in the final chapter.

CHAPTER NINE

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

Diagnosis is not the end, but the beginning of practice.

Martin H. Fischer

9.1. INTRODUCTION

As highlighted by the researcher, the nonprofit sector is becoming an increasingly important sector in a developed economy. The managers of nonprofit organisations (NPOs) need effective performance measurement (PM) systems to manage their organisations and be accountable to their key stakeholders. The existing PM systems for NPOs, including the balanced scorecard (BSC), formed the basis of this study. The study, more precisely the fieldwork, was confined to the nonprofit healthcare sector. The aim of the research was to develop and test a generalisable PM system that could be used by healthcare NPOs to assess their strategic performance. The knowledge gaps identified through the literature review led the researcher to frame four research questions linked to the research aim; the research questions corresponded to five research objectives. The purpose of this chapter is to demonstrate how each research question was answered, what transpired, and how each research objective was achieved, highlighting directions for further research.

The rest of this chapter is organised as follows. Section 9.2 highlights the knowledge gaps, the research questions and the objectives as well as the specific method that was used to achieve the research objectives. Section 9.3 explains what was achieved (and how) under each research objective. Section 9.4 justifies the original contribution (new knowledge) of the study. Finally, section 9.6 highlights future research directions, in the light of the limitations of the researcher's study.

9.2. KNOWLEDGE GAPS, RESEARCH QUESTIONS, RESEARCH OBJECTIVES, AND THE MIXED METHODS APPROACH

The literature review indicated that although the BSC provided a reasonable theoretical basis on nonprofit performance management (see section 3.4.1) relative to other process-based PM models, there are several knowledge gaps that can be filled within a single research study. These gaps relate to both theory and the practice of the BSC in the nonprofit sector.

9.2.1. The Basis for the Research Questions and Objectives

From a theoretical standpoint, the nonprofit BSC was found to be driven by shareholder considerations rather than stakeholder considerations; exclusion of workforce outcomes (particularly people satisfaction) as a performance dimension was found to be the most notable omission on the stakeholder front. The BSC was also found to be lacking clarity on implied causal relationships between its performance dimensions.⁵¹ In addition, the researcher found that the BSC fails to explain how the measures as a whole indicate how an organisation is progressing overall, in implementing its strategies.

Having to formulate strategy maps to implement the BSC, lack of clarity on how performance measures should be developed, limiting the number of PM dimensions to four (or five when the Mission is also included) and limiting the number of performance measures to a smaller number, were found to be the problems from a practitioners' standpoint (details in section 3.6).

Having identified these knowledge gaps, the researcher proposed the following four research questions (for details see section 3.6).

RQ1. What set of PM dimensions constitute an integrated PM system for NPOs?

RQ2. What are the operational definitions of these PM dimensions?

⁵¹ As mentioned in Chapter One, the researcher uses the terms performance dimensions and PM dimensions interchangeably. In the context of this study, these dimensions can also be treated as constructs of the researcher's theoretical model (Figure 9.2).

RQ3. How do the PM dimensions theoretically relate to one another in explaining the achievement of strategic objectives of a NPO?

RQ4: What weighting should managers give to each PM dimension and its subparts in assessing the overall strategic performance of a NPO?

Thus the overall research question of the study was:

What are the performance dimensions that underpin PM in a NPO and how are these dimensions causally related?

The above research questions (these relate to the theory gaps related to the nonprofit BSC) coupled with the practitioner related problems outlined above can be mapped to the research objectives set out by the researcher (section 1.4.2 in Chapter One). For the convenience of the reader, the research objectives are restated as follows.

Objective 1: To determine how the BSC's performance dimensions (domains) should be augmented and operationalised to predict and explain the strategic performance of NPOs.

Objective 2: To develop a scientifically validated performance measurement instrument to test the validity of the BSC's performance dimensions (as augmented).

Objective 3: To test the theoretical causal relationships (i.e. the structural relationships) between BSC measurement domains and interpret the results from a theoretical and practical perspective.

Objective 4: To determine empirical weights for each performance dimension (and its subparts) to enable NPOs to assess their overall strategic performance.

Objective 5: To compare and contrast the researcher's PM framework with the existing PM frameworks—in particular the BSC, which enabled the researcher to initiate her theory development.

The general objective of the study was to develop guidelines for practitioners in NPOs to help them to assess the strategic performance of their organisations.

9.2.2. The Mixed Methods Approach

Figure 9.1 depicts the logical flow of the research. This figure indicates the links between the research questions and research objectives as well as the *major links* between the nine chapters of this thesis. As depicted in Figure 9.1, the researcher used a mixed methods approach (an exploratory sequential design) to achieve the research objectives. The researcher derived a conceptual model from the extant literature to underpin the nonprofit BSC's proposition that the Mission of a NPO drives its strategy to develop and deploy its Strategic Resources (Financial, People and the Infrastructure) to enable its Processes to deliver outcomes to Clients and Donors. In finalising the conceptual model, the researcher treated Clients, Donors and People as the key stakeholders, whose needs need to be satisfied to achieve the Mission of a NPO (the fieldwork validated this). The conceptual model (details in section 4.4 in Chapter Four) formed the theoretical basis of data collection. Data collection consisted of a qualitative phase (a case study) and a quantitative phase.

The qualitative phase involved nine nonprofit healthcare organisations (six from New Zealand and three from Australia), each represented by a senior manager. The researcher collected qualitative data (via a questionnaire followed up by an interview to gain further information) on the constructs underpinning the conceptual model to: (a) provide initial answers to RQ1, (b) answer RQ2 (although the quantitative study in the subsequent phase provided further confirmation), and (c) posit a testable model to answer RQ3. In effect, objectives 1 and 2 were achieved via the analysis of results generated from the qualitative phase (see Figure 9.1).

The quantitative phase involved numerical data collected via a Likert type survey questionnaire administered to a sample of nonprofit healthcare organisations across New Zealand and Australia; the survey questionnaire was developed from the results obtained from the qualitative study; that is, from the operational descriptions of the constructs (PM dimensions) of the researcher's final model. The survey data ($n = 223$) enabled the researcher to test the theoretical models to answer RQ3 and RQ4 as well as

address the remaining research objectives. The next section summarises how each research objective was achieved.

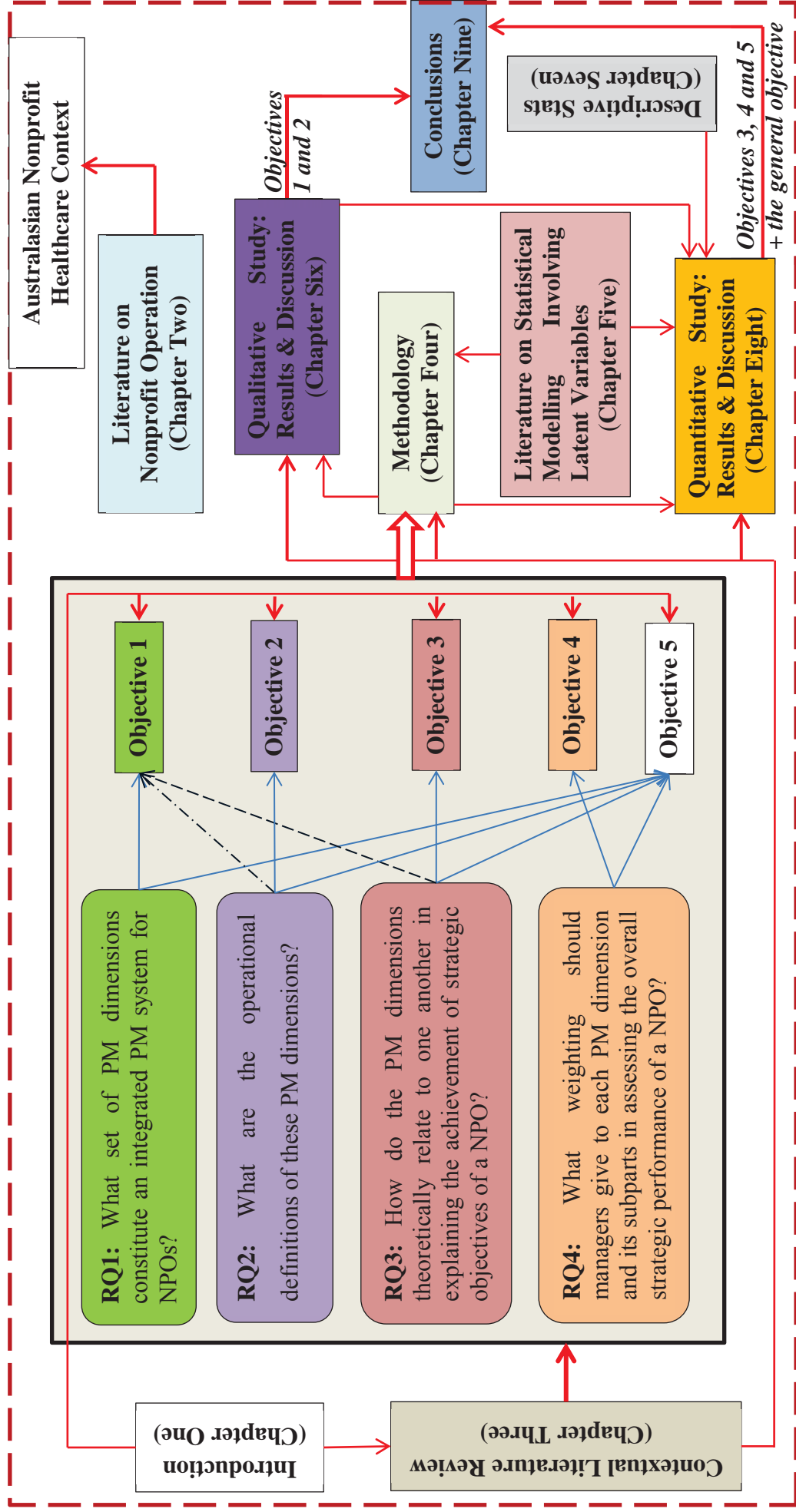


Figure 9.1: The links between the research questions and research objectives as well as the links between the thesis chapters

9.3. FINDINGS FOR EACH RESEARCH OBJECTIVE

The conclusions of the findings under each research objective are as follows.

9.3.1. Findings on Objective 1

Objective 1: To determine how the BSC's performance dimensions (domains) should be augmented and operationalised to predict and explain the strategic performance of NPOs.

Objective 1 was achieved successfully by answering the first research question (RQ1), although findings on the next two research questions (RQ2 and RQ3) were also indirectly used to achieve objective 1 (see Figure 9.1). The conceptual model developed from the literature to underpin the nonprofit BSC consisted of six PM constructs: Mission, Strategy, Economic Health, Organisational Capabilities and People Development, Processes, and Stakeholder Satisfaction. The qualitative data—the written responses provided by the respondents in nine case study NPOs plus the verbal responses provided by them in the interviews—synthesis enabled the researcher to gain a fuller understanding of the phenomenon being studied (understanding the theoretical underpinnings of the nonprofit BSC), as well as operational descriptions of the underlying constructs. The constructs of the refined model developed through the qualitative study (Figure 9.2) are the PM dimensions that constitute an integrated PM system for NPOs.

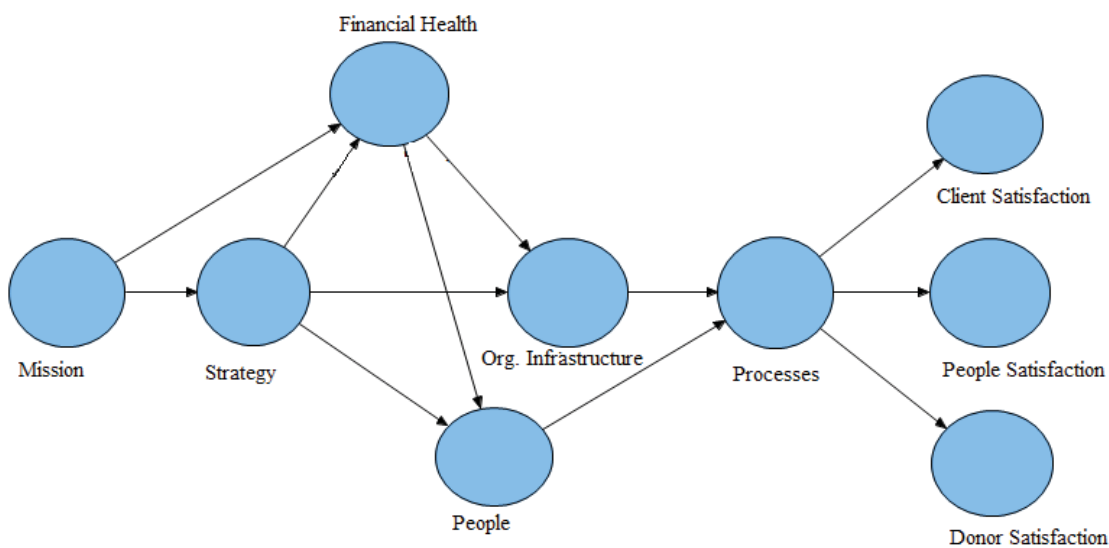


Figure 9.2: The finalised theoretical model on strategic performance

However, the validity of the constructs of the refined model (the final model) was tested in the next phase through the quantitative study using the partial least squares based structural equation modelling (PLSBSEM) approach. The theoretical model (Figure 9.2) was found to be a good fit to the data and the constructs were shown to pass the psychometric tests on construct validity and scale reliability (details in section 8.3 in Chapter Eight). For example, the operational descriptions/definitions of the constructs (performance dimensions) were found to be empirically valid, in terms of the convergent validity and discriminant validity (details in section 8.2 in Chapter Eight). In nontechnical language, this means that the constructs have been operationalised correctly. The implications of these findings were summarised later (section 9.3.3). Based on the above evidence, the researcher argues that *objective 1 has been achieved* by augmenting the nonprofit BSC into a testable theoretical model (Figure 9.2) via the qualitative study and testing the validity of the model via the quantitative study.

9.3.2. Findings on Objective 2

Objective 2: To develop a scientifically validated performance measurement instrument to test the validity of the BSC's performance dimensions (as augmented).

Needless to say, to develop a scientifically validated PM instrument to test the validity of the BSC's augmented PM dimensions (achieving objective 2), it becomes necessary to operationalise the PM dimensions (answering RQ2). As mentioned earlier, the qualitative study enabled the researcher to explore the rich information pertaining to the researcher's final model (Figure 9.2). As mentioned elsewhere (section 4.5.4 in Chapter Four and section 6.3 in Chapter Six), the qualitative data were processed by coding and analysing the data using the software 'Interview Streamliner'. Through this exercise, the researcher was able to identify the related qualitative information on each theme (construct/domain/PM dimension) and subtheme (sub-construct/subdomain/sub-dimension) to isolate the operational descriptions for each construct (PM dimension) in the final model. For the convenience of the reader, the operational descriptions of the constructs provided in Chapter Six are reproduced in Table 9.1.

Table 9.1: Operational Descriptions of Each PM Dimension

Mission
5. A NPO exists to achieve social outcomes; social profit should forefront
6. Improving the experience of the community should be highlighted in the mission
7. Staff development should be highlighted in the mission
8. Need to pursue leadership in what a NPO does to energise staff
Strategy
6. Triple bottom line elements—donors, community, and the environment—should be considered in strategy planning
7. Senior leadership involvement is important not only for strategy planning but also for strategy implementation
8. Achieving social responsibility requires collaboration with key stakeholders, pursuing continuous improvement
9. Organisational structure and resources are considered in strategy implementation
10. In operational planning, work has to be mapped to people to motivate them to perform
People
4. Senior leadership’s capability gives strength to the organisation to drive its processes
5. Decision making should be “business like”; “nonprofit” means non-profit initiatives
6. Creating a high performance work culture is as important to a NPO as for a for-profit organisation
Organisational Infrastructure
3. Information technology (IT) is vital for developing new services and communicating with the stakeholders
4. Creating “brand awareness” and public awareness via the information system is a critical organisational capability
Financial Health
7. Investing in IT/infrastructure strengthens the financial position as this improves

productivity, which in turn attracts funding
8. Cost reduction (direct costs and overheads) improves the financial position
9. Providing regular feedback to sources who provide funding increases the chance of sustaining the funding
10. Relationships with stakeholders who provide funding is important for financial sustainability
11. Innovation initiative development of new services attracts of new funding
Processes
9. Process improvement should be aimed at achieving better outcomes for donors; utilising infrastructure, technology and materials
10. Process improvement should be aimed at achieving better outcomes for clients; safe work processes are part of this
11. Operational strategies are as important as corporate level strategies
12. Support processes are part of continuous improvement
13. New product development should take into account the community and the funder requirements
14. Processes should also include IT/technology development
15. Providing the resources and creating the right organisational structure is vital in strategy implementation
16. The effectiveness of strategies need to be monitored and ineffective strategies need to be replaced
Client Satisfaction
5. Intense service audits usually result in client satisfaction; sometimes donor satisfaction also follows
6. Quality of life of clients' needs to be measured/audited to assess and improve their satisfaction
7. Association to Client feedback on the services results in their satisfaction
8. Client satisfaction needs to be measured regularly
People Satisfaction
7. Paid employees should be remunerated appropriately
8. Training improves employees' satisfaction
9. Volunteer development programmes must not be overlooked
10. Communicating strategic business plans to staff motivate them intrinsically

11. Staff satisfaction need to be measured regularly
12. Recognition of employees and volunteers improves their motivation and satisfaction
Donor Satisfaction
4. Processes designed around client needs usually result in client satisfaction, which ultimately results in donor satisfaction
5. Developing new service models results in donor satisfaction (because this usually results in client satisfaction)
6. Fulfilment of social responsibility requires collaboration with other organisations

The operational descriptions were then transformed into a self-administered survey questionnaire (survey instrument) for quantitative data collection (see Appendix C for the full questionnaire) to test the validity of the performance dimensions (constructs) of the final model. The researcher then went on to further describe the scope of each performance dimension (details in section 6.6 in Chapter Six) in the model and discuss the implications of the findings in the qualitative phase of the study from a theoretical as well as practical standpoint. A very important practical contribution of the qualitative phase of the study is finding a set of critical success factors (CSFs) for each performance dimension; a set of key performance indicators (KPIs) on each CSF was also suggested by the researcher (see Table 6.6 in Chapter Six).

For these reasons, the researcher concludes that RQ2 has been answered and research objective 2 has been successfully achieved.

9.3.3. Findings on Objective 3

Objective 3: To test the theoretical causal relationships (i.e. the structural relationships) between BSC measurement domains and interpret the results from a theoretical and practical perspective.

Objective 3 was achieved successfully by comprehensively answering the third research question (RQ3), using quantitative data (the 223 survey responses). The respondents of the quantitative study were senior managers representing a sample of healthcare NPOs

in Australia and New Zealand (see Chapter Seven for the display of descriptive statistics) and therefore, the data, which are of a strategic planning nature, can be treated as reliable and trustworthy. The statistical analysis also showed that the data are relatively free from common method bias (details in Chapter Eight). The quantitative data tested the following hypotheses.

H1: Mission has a positive effect on Strategy.

H2: Mission has a positive effect on Financial Health.

H3: Strategy has a positive effect on the Financial Health.

H4: Strategy has a positive effect on Organisational Infrastructure.

H5: Financial Health has a positive effect on Organisational Infrastructure.

H6: Strategy has a positive effect on People.

H7: Financial Health has a positive effect on People.

H8: Organisational Infrastructure has a positive effect on Processes.

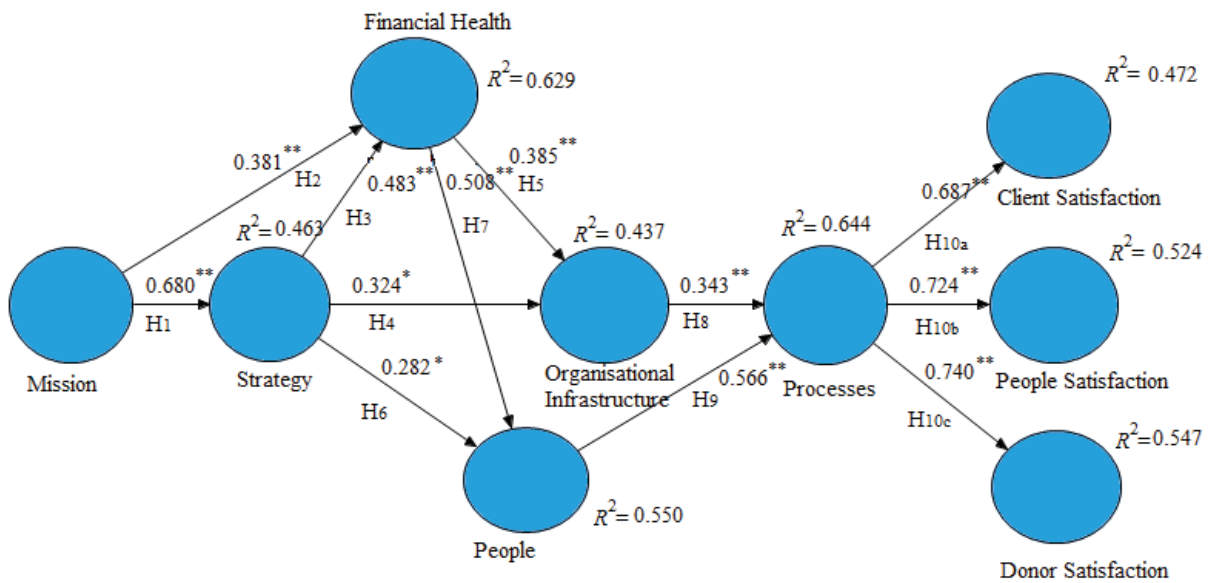
H9: People have a positive effect on Processes.

H10a: Processes have a positive effect on Client Satisfaction.

H10b: Processes have a positive effect on People Satisfaction.

H10c: Processes have a positive effect on Donor Satisfaction.

For the convenience of the reader, the statistical analysis results involving the hypothesised causal paths of the main model are reproduced in Figure 9.3.



Note: **p < 0.001, *p < 0.01

Figure 9.3: The structural relationships between constructs and parameter estimates

As evidenced from the parameter estimates shown in Figure 9.3, all the specified hypotheses are supported by data. In addition, the results also indicate that the hypothesised relationships are relatively strong. For example, H1 is supported by data because the standardised path coefficient (structural regression coefficient) corresponding to H1 (= 0.680) is statistically significant ($p < 0.001$). In addition, H1 implies a relatively strong relationship because 46.3% ($R^2 = 0.463$) of the variability of the Strategy is explained by the Mission.

The researcher highlighted that although the founder of the nonprofit BSC asserted that the Strategy should be quantified and measured and that the Mission of a NPO should be “featured and measured” (Kaplan, 2008, p. 361), the BSC does not provide direct guidance as to how this could be accomplished (details in section 8.4 in Chapter Eight). This is because, as argued by the researcher, the BSC places great emphasis on drawing strategy maps (these maps explain how strategic objectives causally relate to achieve the Mission), which in practice, is not tenable in most situations. The researcher argued that her model directly addresses the difficulties of quantifying and measuring Strategy and Mission because her model features these concepts as measurement domains.

The practical implications of the parameter estimates were also discussed by the researcher in detail (section 8.5 in Chapter Eight). The researcher argued that People are more important for the Processes than the Organisational Infrastructure (a standardised regression coefficient of 0.566 versus a standardised regression coefficient 0.343). On a similar vein, the researcher argued that the Financial Health is more important for People management and development than for Organisational Infrastructure management and development. The researcher also showed that the Processes of a NPO seem to satisfy the three key stakeholders equally because the standardised path coefficients and the R^2 values associated with Client Satisfaction, People Satisfaction and Donor Satisfaction were not practically very different. This also seems to be consistent with *normative stakeholder theory*, which implies that no legitimate stakeholder should receive more priority than any other.

Thus the researcher concludes that her model was found to be acceptable both theoretically (a good fit to the data) and practically (hypothesised relations between constructs) within an Australasian nonprofit healthcare context and that the third research objective were achieved satisfactorily.

9.3.4. Findings on Objective 4

Objective 4: To determine empirical weights for each performance dimension (and its subparts) to enable NPOs to assess their overall strategic performance.

Objective 4 was achieved successfully by answering the fourth research question (RQ4). First, a hierarchical measurement model containing the six constructs (reducing the nine constructs of the main model to six constructs through rationalisation) was posited to represent the Overall Performance of a NPO. The Overall Performance was represented as an index named the Overall Performance Index (OPI) by linearly combining the six predictor constructs via the corresponding standardised regression coefficients (see the λ_i values shown in Figure 9.4). The six constructs (categories) were in turn represented as a linear combination of their constituents (items) through the weight parameters w_{ij} , as shown in Figure 9.4. The unknown parameters (λ_i and w_{ij}) were then estimated via the PLSBSEM technique by fitting the hierarchical model to quantitative data that were used to test the researcher's main model (Figure 9.2). Since the parameter estimates

were found to be statistically significant ($p < 0.001$) and consistent (> 0), the researcher went on to convert the parameter estimates to nominal points by treating the overall performance to be equal to 1000 points. The points (weights) thus estimated for each performance dimension (category) and its subparts (items) are reproduced in Table 9.2 for the convenience of the reader (for details see section 8.7 in Chapter Eight).

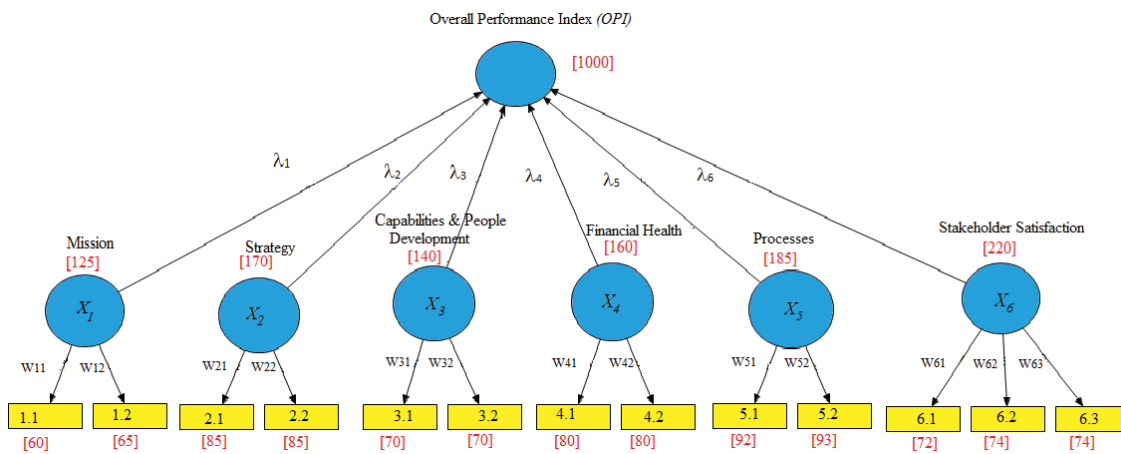


Figure 9.4: The hierarchical model on overall performance

Table 9.2: The Point Scoring System

Measurement Categories	Measurement Items (Subparts of the Construct)	Estimated Points	
		Item	Category
1. Mission	1.1 Envisioning Social Profit	60	125
	1.2 Envisioning Organisational and People Excellence	65	
2. Strategy	2.1 Strategic Planning	85	170
	2.2 Strategy Implementation	85	
3. Organisational Capabilities and People Development	3.1 Organisational Infrastructure	70	140
	3.2 People	70	
4. Financial Health	4.1 Cost Control	80	160
	4.2 Cash Flow Increase	80	
5. Processes	5.1 Continuous Improvement and Knowledge Management	92	185
	5.2 Well Designed, Safe, Efficient and Effective Processes	93	

Measurement Categories	Measurement Items (Subparts of the Construct)	Estimated Points	
		Item	Category
6. Stakeholder Satisfaction	6.1 Client Satisfaction	72	220
	6.2 People Satisfaction	74	
	6.3 Donor Satisfaction	74	
Overall Performance		1000	1000

The researcher argued that her approach of fitting a performance index model to empirical data provides a useful and unbiased performance scoring system that NPOs can use as a dashboard to monitor and improve their performance. In addition, the researcher showed that the government agencies and other major donors who provide substantial funding to NPOs can use the scoring system to monitor the relative performance of their beneficiaries. On the strength of these findings (more details in section 8.7 in Chapter Eight), the researcher argues that RQ4 has been comprehensively answered and thereby, the fourth research objective has been achieved satisfactorily.

9.3.5. Findings on Objective 5

Objective 5: To compare and contrast the researcher’s PM framework with the existing PM frameworks—in particular the BSC, which enabled the researcher to initiate her theory development.

Objective 5 was successfully achieved by the researcher by gaining knowledge in answering the research questions. By synthesising the extant literature and answering the research questions, the researcher was able to gain a full understanding of her model to compare it with the existing PM frameworks, including the BSC (details in section 8.9 in Chapter Eight). A brief summary of the salient features of the researcher’s PM model is given in this section.

In the researcher’s model, the Mission is posited to drive the Strategy and both Mission and Strategy are explicitly featured and measured. The accessory that the BSC uses to feature the Strategy and Mission (and also to ensure that the Mission drives the Strategy) is the strategy map. The researcher’s model avoids the use of strategy maps. The fieldwork as well as the literature indicated that drawing strategy maps is a difficult

proposition for NPOs, and that in many instances, it is untenable to make strategy maps a prerequisite for developing a PM system for NPOs. The researcher addressed this problem by developing a generalisable theory from the nonprofit BSC; in the researcher's causal model (Figure 9.2), nine PM dimensions (constructs) are featured in a causal predictive manner to explain the strategic performance of a NPO, whereas in the BSC, the causal connections between the PM dimensions are (at best) implied. Moreover, through fieldwork, the researcher was able to operationally describe/define each PM dimension comprehensively. In addition, the operational descriptions of each PM dimension were topped up with CSFs (and a sample of KPIs for each CSF) to make the application of the researcher's PM framework more user-friendly.

The researcher's model makes stakeholder satisfaction more explicit; the model posits that Processes result in Client Satisfaction, People Satisfaction, and Donor Satisfaction. Unlike the BSC, the researcher's model makes People Satisfaction (i.e. the satisfaction of employees and volunteers gain from their work) an explicit performance outcome of a NPO. The BSC values the intrinsic motivational needs of the People and that is why Learning and Growth feature as the driver of Processes in the BSC. However, the BSC does not make explicit checks and balances as to what extent People are actually satisfied by their work. One can argue that the BSC does not preclude inclusion of measures on *people-related outcomes*, such as employee turnover and occupational health metrics, under the Learning and Growth dimension or the Internal Business Processes dimension. However, the inclusion or exclusion of such metrics remains entirely at the discretion of the user. Besides, inclusion of metrics on the satisfaction of people under the Learning and Growth dimension or Internal Business Processes dimension ignores the fact that satisfaction results from the organisational activities (Processes) delivering value to the stakeholders. Thus the researcher argues that her model is more aligned to the stakeholder view of corporate governance than the BSC. The BSC has never been a model that fully aligns with *stakeholder theory*. In fact, as shown by the researcher elsewhere, the regular BSC (the for-profit version) is aligned with the *shareholder view* or at best, the *instrumental view of stakeholders*; the nonprofit BSC, of course, downplays the emphasis on financial metrics, as creating long-term shareholder value does not apply in a nonprofit context. However, in augmenting the nonprofit BSC, the researcher ensured that she preserved the spirit of the BSC in not trying to overemphasise stakeholder theory.

The researcher also provides a wider definition on Processes than the BSC's Internal Business Processes dimension. This is because, in addition, the *critical processes* and *innovation processes* (the two subdomains of the Internal Business Processes dimension of the BSC), the researcher treats *continuous improvement* as an explicit subdomain in operationalising the Processes. This brings the researcher's nominal definition and the operational definition of Processes more in line with contemporary quality management theory.

The researcher observes that organisational excellence models such as the Baldrige Framework and the EFQM Excellence Model take a similar stance on Stakeholders and Processes as the researcher; however, these models are promulgated by their custodians as universal solutions for organisations—a single theory (model) on performance excellence and a single scoring system (weighting of performance dimensions and sub-dimensions in determining the overall performance), no matter what the organisation is (e.g. for-profit, government, or nonprofit). The researcher sees performance measurement in this fashion as treating PM models as 'Lycra'. The researcher showed that for-profit performance management and nonprofit performance management diverge in some respects (e.g. what drives the system, what sets the strategies, and the differences in the governance structures).

In summary, the researcher's PM model shows clear links between the Mission, Strategy and the remaining elements of the system (inputs, processes, and outputs), which is not clearly delineated in the BSC. The researcher's PM model also plays the role of an organisational performance theory, because it explains how the organisation adds value to the community or the targeted group, since it measures stakeholder satisfaction (hence mission achievement) by aligning the organisational Mission, Strategy, Resources (Financial Health, People, and the Organisational Infrastructure) and Processes in a causal predictive manner; equally importantly, the researcher validated her model empirically, for the populations to which then model is meant to be generalisable (i.e. nonprofit healthcare organisations). Given these findings, the researcher concludes that the fifth research objective has been achieved.

9.3.6. Findings on the General Objective

General Objective: To develop guidelines for practitioners in NPOs to help them to assess the strategic performance of their organisations.

Finally, the general research objective was achieved by the researcher by being able to accomplish two primary tasks: (a) by developing CSFs for each PM dimension of the researcher's PM model to help managers develop KPIs for their own organisation, and (b) by prescribing how the researcher's quantitative survey questionnaire and the scoring system can be used for strategic decision making.

In Chapter Six (see section 6.7.2 for details), the researcher showed that each operational description that the researcher provided for each PM dimension of her model, neatly translates into a CSF (thus 40 CSFs for the 40 operational descriptions belonging to the 9 PM dimensions); this made generation of KPIs more straightforward than through strategy maps, as prescribed in the BSC. In addition, the researcher mentioned that the exemplar KPIs shown can also be used by managers, although the researcher highlighted that the CSFs are more robust than the KPIs and therefore, it is desirable for managers to develop their own KPIs from the CSFs given by the researcher.

In Chapter Eight (see section 8.8.3 for details), the researcher showed how a NPO could conduct a self-assessment to assess its strategic performance, using the researcher's quantitative questionnaire and the scoring system (see Table 9.2). The researcher recommended several steps for nonprofit practitioners to follow to obtain a representative score for each PM item in each PM dimension to assess the overall strategic performance of their organisation. The researcher cautioned that an organisation should not spontaneously react to a low item score as long as it has scored well in the other items in the same category. This is because a low item score can sometimes be a consequence of incomplete information. The researcher also highlighted that because the scoring system (and thus by default, self-assessment guidelines) has been derived from the researcher's hierarchical model on overall performance (Figure 9.4), which in turn is linked to the researcher's main model (Figure 9.2), the self-assessment guidelines and the scoring system should always be used in

conjunction with the researcher's main model. It is the main model that predicts and explains how strategic objectives are achieved.

9.4. JUSTIFYING THE ORIGINAL CONTRIBUTION

The theoretical and practical contribution of the study was stated at the beginning of this thesis (section 1.6 in Chapter One). In the current section, the researcher justifies the statements made earlier on the original contribution.

9.4.1. The Theoretical Contribution

The primary theoretical contribution of the study comes from the augmentation of the theoretical underpinnings of the nonprofit BSC and testing the researcher's model empirically.

9.4.1.1. Theory Development

As mentioned elsewhere, the nonprofit BSC posits that the Mission of a NPO drives its Strategy to develop and deploy its strategic resources (Financial, People and the Infrastructure) to enable its Processes to deliver outcomes to Clients and Donors. Although the nonprofit BSC contends that the Mission and Strategy should feature in the BSC and that these should be quantified, neither the Mission nor the Strategy appears as performance dimensions (constructs) in the scorecard. In the nonprofit BSC, the Strategy is represented by the *strategic objectives* and the notion that the Mission should drive the Strategy is embedded in the *strategy map*, which, in nonprofit management, is a difficult causal diagram to construct (without hiring external consultants). The researcher therefore represented the underlying theory of the nonprofit BSC by incorporating Mission and Strategy as constructs (performance dimensions) to circumvent the need for strategy maps. These are not the only new constructs that the researcher added to improve the clarity of the nonprofit BSC.

The Learning and Growth dimension of the nonprofit BSC (the driver of the Process dimension) was represented by the researcher as two separate dimensions—Organisational Infrastructure and People—and the Financial Health was represented as a dimension to reflect managerial action (cash inflow improvement interventions as well as cost control interventions) that is required for the development and management of

People and the Organisational Infrastructure. Stated alternatively, Organisational Infrastructure and People serve as inputs to the Processes and the Financial Health serves as the input to Organisational Infrastructure and People. On the other hand, the researcher represented Donor Satisfaction as an outcome of the Processes. From the fieldwork, the researcher identified three organisational initiatives that create an impact on the Donors (see the three operational descriptions of Donor Satisfaction shown in Table 9.1). A user of the nonprofit BSC may consider the Financial perspective (dimension) of the BSC as a mixture of what the researcher defines as the Financial Health dimension and Donor Satisfaction dimension (two related but separate dimensions). Finally, the researcher highlights that Processes should deliver not only outcomes for the Donors and the Clients but also or the People of the NPO, if they are to achieve the Mission. The nonprofit BSC considers only the Donor outcomes (how the organisation looks to its financial donors) and Client outcomes (how the organisation looks to its clients) explicitly. Development of the operational descriptions/definitions for the constructs (performance dimensions) of the researcher's main theoretical model (Figure 9.2) *for the nonprofit healthcare sector*, and specifying the causal hypotheses involving the performance dimensions are all part of new knowledge.

Development of the hierarchical model (sometimes referred to as the ancillary model) to capture the overall performance of a NPO and to explain how each performance dimension and its subparts contribute to overall performance is also part of researcher's theory (new knowledge) development.

9.4.1.2. Theory Testing

Testing the researcher's models (the main model and the ancillary model) through large sample quantitative data (from the nonprofit healthcare sector) is the other important theoretical contribution made by the researcher. This is unique because the study is the first of its kind that attempts to examine the statistical generalisability of any PM framework across the nonprofit sector in Australasia. The researcher's causal hypotheses (the main model) were well supported by the data and the unknown statistical parameters estimated via the researcher's ancillary model (Figure 9.4) enabled the researcher to demonstrate to the academia that her ancillary model is tenable.

In concluding section 9.4.1, the researcher justifies that her work contributes to the growing body of literature on the adaptability of the BSC and helps academia to expand their current knowledge on the theoretical basis of the BSC dimensions. The academic publications that have resulted from the study so far, have been listed elsewhere.

9.4.2. The Practical Contribution

The study resulted in multiple practical contributions. Firstly, the researcher's PM framework (the main theoretical model) can be used as a tool by the practitioners to understand or explain successful strategy implementation versus unsuccessful strategy implementation. This is because the researcher's PM framework explains how strategic outcomes are caused. Secondly, the PM instrument (the quantitative questionnaire) developed by the researcher to test her models can be used by the practitioners to self-assess how their organisation is performing in each PM dimension. The numerical information on overall performance generated from the PM instrument (i.e. performance feedback) in conjunction with the scoring system developed by the researcher help managers to take corrective action when needed. In addition, unlike the nonprofit BSC, because the Mission and Strategy are directly measured in the researcher's models, nonprofit managers can react to strategy ineffectiveness more swiftly than otherwise.

The final practical contribution comes from the guidelines provided by the researcher on how to use the researcher's findings (e.g. the CSFs, the quantitative questionnaire and the scoring system) to improve the strategic performance of a (healthcare) NPO.

9.5. METHODOLOGICAL LESSONS LEARNT

The study resulted in exposing the researcher to several important methodological aspects, which can be regarded as methodological lessons learnt from the research. For the benefit of researchers conducting similar research, these are as follows.

- In survey research, timing of launching the questionnaire is very important.
- Although it is desirable to obtain responses from a wider net of respondents of an organisation—top level managers, middle level managers, lower level

managers and the frontline workers— awareness (or the lack of it) of strategic issues limited the data collection to managers at the very highest level.

- It is necessary to make a compromise between what is ideal and what is possible (the issue of using the Likert scale in PLSBSEM is a classic example).
- Top managers in the nonprofit sector are very responsive/supportive of research pertaining their organisations; they do not seem to resist interviews.
- The researcher's background has a significant impact on selecting the research paradigm and the relevant methodology.

Some of the above points are elaborated briefly for clarity.

The researcher realised that collecting large sample data in the December to February period is not ideal in Australia and New Zealand. This is due to the summer vacation in the Southern Hemisphere during this period. Although the researcher did not postpone data collection to March and beyond due to tighter timeframes, she had to nonetheless collect additional data subsequently (see Figure 7.2) to make the sample size adequate. Thus, in the end, there was no real advantage in launching the questionnaire in early summer (late November/early December). If possible, large sample data collection should be planned in the March to August (no later than September) period in the Southern Hemisphere.

The researcher also learnt that it is necessary to make optimum (the best) decisions in selecting the right research technique/s. For example, the Likert scale used in the questionnaire is an ordinal scale (a Likert scale is always ordinal) but the researcher had to treat that scale as 'interval' as other researchers do, in order to conduct PLSBSEM; since there are no known PLSBSEM software packages that specifically handle ordinal data, a justifiable approximation was needed. Under these circumstances, as prescribed in the literature, ordering of response choices was carefully done (e.g. selecting the meaningful midpoint "neither agree nor disagree") to make the interval approximation tenable.

Prior to undertaking the doctoral study, the researcher worked as a lecturer in engineering mathematics and statistics in a leading university in Sri Lanka. In addition, her prior postgraduate training was on financial mathematics. The researcher realised that this background had a significant impact on shaping the research inquiry. In retrospect, what the researcher has done was what positivism always advocates: operationalise the concepts, formulate the hypotheses, and test the hypotheses (albeit within a mixed platform).

9.6. LIMITATIONS AND FURTHER RESEARCH DIRECTIONS

Any good theory—even Newtonian physics—operates within a specific scope and a boundary within which the theory is supposed to be generalisable; the characteristics that define the scope and the boundaries of a theoretical contribution is known as the delimitations. In addition, any study is associated with certain uncontrollable factors that impose limitations to the study (some uncontrollable factors are negated through assumptions). Several future research directions can be recognised, based on the findings, limitations and delimitations of the research.

The generalisability of the findings of the study is restricted to the Australasian nonprofit healthcare sector, because the researcher collected qualitative data (for operationalisation of the constructs) and quantitative data (for theory testing and parameter estimation) from that sector only. Also, the researcher assumed that there are no situational factors (e.g. the between-country/cultural differences, sub-sector differences, and the organisation size differences within the nonprofit healthcare sector) that affect the hypothesised causal relationships (Figure 9.3) and the points scoring system (Table 9.2) developed by the researcher. In reality, there may be situational factors that affect (moderate) the researcher's findings. Another limitation of the study is that the researchers' PM framework and the PM instrument could not be tested for practical usability/validity in a longitudinal type of a study. Finally, the study does not examine in any great detail - how the findings can be used in public sector performance management. The future research directions therefore are as follows.

9.6.1. Enhancing the Generalisability

The nonprofit sector is a very diverse/heterogeneous sector, which, for the purpose of classification, has been divided into twelve main activity groups (the international classification of NPOs), as shown in Chapter Two. There could be differences within an activity group and, more so, between the activity groups, in regards to performance management. Out of the twelve main activity groups, the researcher selected the activity group “health” to set the context of the study. To enhance the generalisability of the findings and to identify any sectorial differences (i.e. between activity group differences) on performance management, future research may be undertaken to fit the researcher’s models (the main model and the ancillary model) to data collected from other activity groups and other countries within the Anglo-Saxon cultural cluster.

9.6.2. Studying Potential Moderating Effects

The researcher’s model does not consider any moderating factors that may affect the theoretical relationships specified in her models. There could be situational factors that could moderate the hypothesised relationships. Thus, further research may involve identifying certain factors and their effects on the hypothesised relationships. The moderating factors could be, as mentioned earlier, the size of the organisation (e.g. medium versus large), the level of nongovernmental financial income (low versus high), the proportion of paid employees (or volunteers), and the type of service (activity group/subgroup could be used as a proxy). The moderation factors should be selected cautiously because the model becomes more and more complex (e.g. theory testing wise) as one adds each moderating factor.

9.6.3. A Longitudinal Study for Enhancing the Practical Usability

As mentioned earlier, due to resource constraints (mainly the time available), the researchers’ project deliverables (PM framework, the PM instrument, and the scoring system) could not be implemented in NPOs to test their practical usability. Thus future research could be directed towards understanding what gains NPOs achieve by applying the researcher’s practical contributions to achieve mission-related outcomes. Performance measurement and management is a social process, and therefore the attitudes and values of the People, Clients and the wider community can affect the effectiveness of a PM system. Practical implementation of the researcher’s findings in a

panel type study (a longitudinal study), helps other researchers to identify how best the findings of this study work in practice, and what else is needed to make the findings more useful to the practitioner.

9.6.4. Adapting to the Public Sector

Both the public sector and the nonprofit sector provide social goods. Typically, the PM fraternity treat performance management in both sectors as being similar. However, there are differences between the two sectors and this may mean that the findings of this study could be less useful to the public sector. For example, the concept “charitable giving” is not applicable to the public sector (under normal circumstances) and the streams of revenue creation and the profile of people in the public sector could be quite different from the nonprofit sector. Therefore, future studies may be directed towards identifying the applicability of the researcher’s key findings to the public sector.

REFERENCES

- Adebanjo, D., Abbas, A., & Mann, R. (2010). An investigation of the adoption and implementation of benchmarking. *International Journal of Operations and Production Management*, 30(11), 1140-1169. doi: 10.1108/01443571011087369
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317-332.
- Akingbola, K. (2013). A model of strategic nonprofit human resource management. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 24(1), 214-240.
- Al-Tabbaa, O., Gadd, K., & Samuel, A. (2013). Excellence models in the non-profit context: Strategies for continuous improvement. *International Journal of Quality & Reliability Management*, 30(5), 590-612. doi: 10.1108/02656711311315521
- Albareda, L. (2008). Corporate responsibility, governance and accountability: from self-regulation to co-regulation. *Corporate Governance: The international journal of business in society*, 8(4), 430-439.
- Alexander, J., Brudney, J. L., & Yang, K. (2010). Introduction to the symposium: Accountability and performance measurement: The evolving role of nonprofits in the hollow state. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 565-570.
- Allison, M., & Kaye, J. (2001). Characteristics of Nonprofit Organizations-Implications for Consultation. *Compass Point Nonprofit Services*.
- Anand, M., Sahay, B., & Saha, S. (2005). Balanced scorecard in Indian companies. *Vikalpa*, 30(2), 11-25.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423. doi: 10.1037/0033-2909.103.3.411
- Andreaus, M., & Costa, E. (2014). Toward an integrated accountability model for nonprofit organizations. In E. Costa, L. D. Parker, & M. Andreaus (Eds.), *Accountability and Social Accounting for Social and Non-Profit Organizations* (Vol. 17, pp. 153-176): Emerald Group Publishing Limited.
- Anon. (2008). Can I apply Balanced Scorecard in a non-profit organisation?. Maidenhead, Berkshire: 2GC Active Mgt.
- Anon. (2015). The Performance Prism. Retrieved 22.01.2016, from <http://www.accaglobal.com/ca/en/student/exam-support-resources/professional-exams-study-resources/p5/technical-articles.html>
- Anthony, R. N., & Govindarajan, V. (2007). *Management control systems* (12th ed.). Boston: McGraw-Hill/Irwin.
- Antonsen, Y. (2014). The downside of the Balanced Scorecard: A case study from Norway. *Scandinavian Journal of Management*, 30(1), 40-50.
- Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*: Springer Science & Business Media.
- Armstrong, M. (2014). *Armstrong's handbook of human resource management practice* (13th ed.). Philadelphia, PA: Kogan Page Ltd.
- Atkinson, H. (2006). Strategy implementation: A role for the balanced scorecard? *Management Decision*, 44(10), 1441-1460.
- Austin, J. T., Boyle, K. A., & Lualhati, J. C. (1998). Statistical conclusion validity for organizational science researchers: A review. *Organizational research methods*, 1(2), 164-208.

- Austin, M. J. (2003). The changing relationship between nonprofit organizations and public social service agencies in the era of welfare reform. *Nonprofit and Voluntary Sector Quarterly*, 32(1), 97-114.
- Australian Bureau of Statistics (2015). Australian National Accounts: Non-Profit Institutions Satellite Account, 2012-13. Retrieved June 26, 2016, from [http://www.abs.gov.au/ausstats/abs@.nsf/Products/8106.0~2006-07+\(Re-Issue\)~Main+Features~Overview?OpenDocument](http://www.abs.gov.au/ausstats/abs@.nsf/Products/8106.0~2006-07+(Re-Issue)~Main+Features~Overview?OpenDocument)
- Avery, G., Everett, A., Finkelde, A., & Wallace, K. (1999). Emerging trends in Australian and New Zealand management development practices in the twenty-first century. *Journal of Management Development*, 18(1), 94-108.
- Awadallah, E. A., & Allam, A. (2015). A Critique of the Balanced Scorecard as a Performance Measurement Tool. *International Journal of Business and Social Science*, 6(7), 9.
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative science quarterly*, 421-458.
- Bakewell, O., & Garbutt, A. (2005). The use and abuse of the logical framework approach *Swedish International Development Cooperation Agency* (pp. 27): International NGO Training and Research Centre.
- Balser, D., & McClusky, J. (2005). Managing stakeholder relationships and nonprofit organization effectiveness. *Nonprofit Management and Leadership*, 15(3), 295-315.
- Bang, H., & Ross, S. D. (2009). Volunteer motivation and satisfaction. *Journal of Venue and Event Management*, 1(1), 61-77.
- Banker, R. D., Chang, H., & Pizzini, M. (2011). The judgmental effects of strategy maps in balanced scorecard performance evaluations. *International Journal of Accounting Information Systems*, 12(4), 259-279.
- Barczak, G., Kahn, K. B., & Moss, R. (2006). An exploratory investigation of NPD practices in nonprofit organizations. *Journal of Product Innovation Management*, 23(6), 512-527.
- Barrett, H., Balloun, J. L., & Weinstein, A. (2012). Creative climate: A critical success factor for 21st century organisations. *International Journal of Business Innovation and Research*, 6(2), 202-219.
- Bass, B. M. (1990). From Transactional to Transformational Leadership: Learning to Share the Vision. *Organizational Dynamics*, 18(3), 19-31.
- Benjamin, L. M. (2013). The potential of outcome measurement for strengthening nonprofits' accountability to beneficiaries. *Nonprofit and Voluntary Sector Quarterly*, 42(6), 1224-1244. doi: 10.1177/0899764012454684
- Bent-Goodley, T. B. (2001). Eradicating Domestic Violence in the African American Community: A Literature Review and Action Agenda. *Trauma, Violence, & Abuse*, 2(4), 316-330.
- Bisbe, J., & Barrubés, J. (2012). The Balanced Scorecard as a management tool for assessing and monitoring strategy implementation in health care organizations. *Revista Espanola de Cardiologia*, 65(10), 919-927. doi: 10.1016/j.rec.2012.05.011
- Bitsch, V. (2005). Qualitative research: A grounded theory example and evaluation criteria. *Journal of Agribusiness*, 23(1), 75-91.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual review of psychology*, 53(1), 605-634.

- Bollen, K. A., & Hoyle, R. H. (2012). Latent Variables in Structural Equation Models. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 56-67). New York: Guilford Press.
- Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305-314.
- Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog*, 139-168.
- Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, *50*(2), 1-5.
- Bou-Llusar, J. C., Escrig-Tena, A. B., Roca-Puig, V., & Beltrán-Martín, I. (2009). An empirical assessment of the EFQM Excellence Model: Evaluation as a TQM framework relative to the MBNQA Model. *Journal of Operations Management*, *27*(1), 1-22. doi: 10.1016/j.jom.2008.04.001
- Bouckaert, L., & Vandenhove, J. (1998). Business Ethics and the Management of Non-Profit Institutions. *Journal of Business Ethics*, *17*(9/10), 1073-1081.
- Bourne, M., Kennerley, M., & Franco-Santos, M. (2005). Managing through measures: A study of impact on performance. *Journal of Manufacturing Technology Management*, *16*(4), 373-395. doi: 10.1108/17410380510594480
- Bourque, L. B. (2003). *How to Conduct Self-Administered and Mail Surveys*. (2 ed.). CA: Sage Publications.
- Bowman, W. (2011). *Finance fundamentals for nonprofits: Building capacity and sustainability*. Hoboken, N.J.: John Wiley & Sons.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research : A practical guide for beginners*. London: Sage
- Brennan, N. M., & Solomon, J. (2008). Corporate governance, accountability and mechanisms of accountability: An overview. *Accounting, Auditing & Accountability Journal*, *21*(7), 885-906.
- Brigham, E. F., & Pappas, J. L. (1976). *Managerial Economics* (2nd ed.). Hinsdale, IL: Dryden Press.
- Britz, G., Emerling, D., Hare, L., Hoerl, R., & Shade, J. (1997). How to teach others to apply statistical thinking. *Quality Progress*, *30*(6), 67.
- Brody, E. (2006). The legal framework for Nonprofit Organizations. In W. W. Powell & R. Steinberg (Eds.), *The nonprofit sector: A research handbook* (pp. 243-266): Yale University Press.
- Brown, M. G. (2013). *Baldrige Award Winning Quality: How to Interpret the Baldrige Criteria for Performance Excellence*. NW: CRC Press.
- Brown, W. A. (2005). Exploring the association between board and organizational performance in nonprofit organizations. *Nonprofit Management and Leadership*, *15*(3), 317-339.
- Bryman, A. (2012). *Social research methods* (4th ed.). Oxford: Oxford University Press.
- Bryson, J. M. (2011). *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement* (Vol. 1). San Francisco, California: Jossey-Bass.
- Bryson, J. M., & Alston, F. K. (2004). *Creating and implementing your strategic plan : A workbook for public and nonprofit organizations* (2nd ed.). San Francisco, Calif. : Jossey-Bass.

- Bussell, H., & Forbes, D. (2002). Understanding the volunteer market: The what, where, who and why of volunteering. *International journal of nonprofit and voluntary sector marketing*, 7(3), 244-257.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. NY: Routledge.
- Calvo-Mora, A., Leal, A., & Roldán, J. L. (2005). Relationships between the EFQM model criteria: A study in Spanish universities. *Total Quality Management & Business Excellence*, 16(6), 741-770. doi: 10.1080/14783360500077708
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research on teaching*: American Educational Research Association.
- Candler, G., & Dumont, G. (2010). A non-profit accountability framework. *Canadian Public Administration*, 53(2), 259-279.
- Caracelli, V. J., & Riggin, L. J. C. (1994). Mixed-method evaluation: Developing quality criteria through concept mapping: Mixed-Method Collaboration. *Evaluation Practice*, 15(2), 139-152. doi: 10.1016/0886-1633(94)90005-1
- Carlucci, D., & Schiuma, G. (2007). Knowledge assets value creation map. Assessing knowledge assets value drivers using AHP. *Expert Systems With Applications*, 32, 814-821. doi: 10.1016/j.eswa.2006.01.046
- Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *American Review of Public Administration*, 39(4), 374-390. doi: 10.1177/0275074008320190
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models causes, consequences, and strategies. *Sociological Methods & Research*, 29(4), 468-508.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern methods for business research*, 295(2), 295-336.
- Chin, W. W. (1999). How to Write Up and Report PLS Analyses. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 655-690): Springer Handbooks of Computational Statistics Series.
- Chin, W. W., & Gopal, A. (1995). Adoption intention in GSS: Relative importance of beliefs. *ACM SigMIS Database*, 26(2-3), 42-64.
- Chin, W. W., & Newsted, P. R. (1999). Structural equation modeling analysis with small samples using partial least squares.
- Chisolm, L. B. (1995). Accountability of nonprofit organizations and those who control them: The legal framework. *Nonprofit Management and Leadership*, 6(2), 141-156.
- Ciasullo, M. V., & Troisi, O. (2013). Sustainable value creation in SMEs: A case study. *TQM Journal*, 25(1), 44-61. doi: 10.1108/17542731311286423
- Clark, C., Rosenzweig, W., Long, D., & Olsen, S. (2004). Double Bottom Line Project Report: Assessing Social Impact in Double Line Ventures, Methods Catalog, Columbia Business School.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Cohen, R. J., & Swerdlik, M. E. (2002). *Psychological testing and assessment: An introduction to tests and measurement* (5th ed.). London: McGraw-Hill.

- Collier, D. A., & Evans, J. R. (2015). *Operations Management* (5th ed.). Boston, USA: Cengage Learning.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research, 12*(3), 229.
- Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology, 25*(3), 325-334.
- Cook, T. D., Campbell, D. T., & Donald, T. (1979). *Quasi-experimentation: Design and analysis for field settings* (Vol. 351). Boston: Houghton Mifflin.
- Cooper, D. R., & Schindler, P. S. (2014). *Business research methods* (12 edition ed.). New York, NY: McGraw-Hill/Irwin.
- Cordery, C. J., & Baskerville, R. F. (2007). Charity financial reporting regulation: A comparative study of the UK and New Zealand. *Accounting History, 12*(1), 7-27.
- Cornwall, A., Lucas, H., & Pasteur, K. (2000). Accountability through participation: Developing workable partnership models in the health sector. *IDS Bulletin, 31*(1), 1-13.
- Costa, E., Ramus, T., & Andreaus, M. (2011). Accountability as a Managerial Tool in Non-Profit Organizations: Evidence from Italian CSVs. *Voluntas: International Journal of Voluntary & Nonprofit Organizations, 22*(3), 470-493. doi: 10.1007/s11266-011-9183-7
- Crampton, P., Lay-Yee, R., & Davis, P. (2004). *Primary Health Care in Community-governed Non Profits: The Work of Doctors and Nurses*: Ministry of Health.
- Creswell, J. W. (2010). Mapping the Developing Landscape of Mixed Methods Research. In A. Tashakkori & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 45-68). Los Angeles: Sage.
- Creswell, J. W. (2014). *Research Design : Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). Thousand Oaks: Sage.
- Creswell, J. W., & Clark, V. L. P. (2011). *Designing and conducting mixed methods research* (2 ed.). LA: Sage
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297.
- Crook, C., & Garrat, D. (2011). The Positivist Paradigm in Contemporary Social Research: The Interface of Psychology, Method and Sociocultural Theory. In B. Somekh & C. Lewin (Eds.), *Theory and methods in social research* (2nd ed., pp. 212 - 219). London Sage.
- Cross, K. F., & Lynch, R. L. (1988). The “SMART” way to define and sustain success. *National Productivity Review, 8*(1), 23-33.
- Cutt, J. (1982). Accountability, efficiency, and the “bottom line” in non-profit organizations. *Canadian Public Administration, 25*(3), 311-331.
- Daft, R. L., & Weick, K. E. (1984). Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review, 9*(2), 284-295. doi: 10.5465/AMR.1984.4277657
- Darke, P., Shanks, G., & Broadbent, M. (1998). Successfully completing case study research: Combining rigour, relevance and pragmatism. *Information Systems Journal, 8*(4), 273-289. doi: 10.1046/j.1365-2575.1998.00040.x
- Davies, B., & Knapp, M. (1994). Improving Equity and Efficiency in British Community Care. *Social Policy & Administration, 28*(3), 263.

- Dawson, T. (1999). Relating variance partitioning in measurement analyses to the exact same process in substantive analyses. *Advances in social science methodology*, 5, 101-110.
- Dechow, N. (2012). The balanced scorecard: Subjects, concept and objects-a commentary. *Journal of Accounting & Organizational Change*, 8(4), 511-527.
- Dees, J. G., Anderson, B. B., & Wei-Skillern, J. (2004). Scaling social impact. *Stanford social innovation review*, 1(4), 24-32.
- Dellinger, A. B., & Leech, N. L. (2007). Toward a Unified Validation Framework in Mixed Methods Research. *Journal of Mixed Methods Research*, 1(4), 309-332. doi: 10.1177/1558689807306147
- Deming, W. E. (1994). The new economics for industry, government, education. *MIT Center for Advanced Engineering Study, Cambridge, MA*.
- Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research* (4th ed.). USA: Sage.
- Dey, C. (2002). Methodological issues: The use of critical ethnography as an active research methodology. *Accounting, Auditing & Accountability Journal*, 15(1), 106.
- Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, 61(12), 1203-1218.
- Dolnicar, S., Irvine, H., & Lazarevski, K. (2008). Mission or money? Competitive challenges facing public sector nonprofit organisations in an institutionalised environment. *International journal of nonprofit and voluntary sector marketing*, 13(2), 107-117.
- Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review*(1), 65.
- Dubé, L., & Paré, G. (2003). Rigor in information systems positivist case research: Current practices, trends, and recommendations. *Mis Quarterly*, 597-636.
- Dubin, R. (1978). *Theory development*. NY: Free Press.
- Duque-Zuluaga, L. C., & Schneider, U. (2008). Market orientation and organizational performance in the nonprofit context: Exploring both concepts and the relationship between them. *Journal of Nonprofit and Public Sector Marketing*, 19(2), 25-47.
- Easterby-Smith, M., Thorpe, R., & Jackson, P. (2012). *Management research* (4th ed.). London: Sage.
- Ebrahim, A. (2003). Accountability in practice: Mechanisms for NGOs. *World development*, 31(5), 813-829.
- Ebrahim, A., Battilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. *Research in Organizational Behavior*, 34, 81-100.
- Eden, C., & Ackermann, F. (2013). *Making strategy: The journey of strategic management*. London: Sage.
- Edwards, M., & Hulme, D. (1996). Too close for comfort? The impact of official aid on nongovernmental organizations. *World development*, 24(6), 961-973.
- EFQM. (2013). The EFQM Excellence Model. 2013(1 February). Retrieved from: <http://www.efqm.org/the-efqm-excellence-model>
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1-26.
- Einfeld, M. (2001). Opening Address, Council of Social Services of NSW Community Summit on Disability Advocacy and Information Services. *New South Wales Council of Social Service (NCOSS)*.

- Elias, A. A. (2015). Analysing the stakes of stakeholders in research and development project management: A systems approach. *R&D Management*.
- Elias, A. A., Cavana, R. Y., & Jackson, L. S. (2002). Stakeholder analysis for R&D project management. *R&D Management*, 32(4), 301-310.
- Emerson, J. (2003). The Blended Value Proposition: Integrating Social and Financial Returns. *California management review*, 45(4), 35-51.
- Emerson, J., Wachowicz, J., & Chun, S. (2000). Social return on investment: Exploring aspects of value creation in the nonprofit sector. *The Box Set: Social Purpose Enterprises and Venture Philanthropy in the New Millennium* (Vol. 2, pp. 130-173). CA: Roberts Foundation.
- Erden, Z., Von Krogh, G., & Nonaka, I. (2008). The quality of group tacit knowledge. *The Journal of Strategic Information Systems*, 17(1), 4-18.
- Eskildsen, J. K., & Dahlgaard, J. J. (2000). A causal model for employee satisfaction. *Total Quality Management*, 11(8), 1081-1094.
- Evans, J. R. (2004). An exploratory study of performance measurement systems and relationships with performance results. *Journal of Operations Management*, 22(3), 219-232.
- Evans, J. R., & Mai, F. (2014). Insights From the Baldrige Award Item-Level Blinded Applicant Scoring Data. *The Quality Management Journal*, 21(2), 45.
- Felzensztein, C., Stringer, C., Benson-Rea, M., & Freeman, S. (2014). International marketing strategies in industrial clusters: Insights from the Southern Hemisphere. *Journal of Business Research*, 67(5), 837-846.
- Fink, A. (2013). *How to conduct surveys : A step-by-step guide* (5 ed.). Calif.: Sage.
- Fink, A., & Litwin, M. S. (2003). *How to assess and interpret survey psychometrics* (Vol. 8): Sage.
- Flack, T., & Ryan, C. (2005). Financial reporting by Australian nonprofit organisations: Dilemmas posed by government funders. *Australian Journal of Public Administration*, 64(3), 69-77.
- Flick, U. (2014). *An introduction to qualitative research* (5th ed.). Los Angeles: Sage.
- Flynn, B. B., & Saladin, B. (2001). Further evidence on the validity of the theoretical models underlying the Baldrige criteria. *Journal of Operations Management*, 19(6), 617-652.
- Folland, S., Goodman, A. C., & Stano, M. (2007). *The economics of health and health care* (Vol. 6). New Jersey: Pearson Prentice Hall.
- Forbes, D. P. (1998). Measuring the unmeasurable: Empirical studies of nonprofit organization effectiveness from 1977 to 1997. *Nonprofit & Voluntary Sector Quarterly*, 27(2), 183.
- Forcada, M. Á. H., Ramis-Pujol, J., & Cusumano, J. L. G. (2008). Comparative Analysis of Nine Performance Measurement Systems. In K. J. F. Foley & H. Philippe (Eds.), *The Theories And Practices of Organisation Excellence: New Perspectives* (pp. 359-388). NSW, Australia: SAI Global Limited.
- Fornell, C. (1987). A second generation of multivariate analysis: Classification of methods and implications for marketing research. In M. J. Houston (Ed.), *Review of marketing* (pp. 407-450). Chicago: American Marketing Association.
- Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing research*, 440-452.
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research* 18(1), 39-50.

- Fowler, J. F. J. (2013). *Survey research methods* (5 ed.). USA: Sage.
- Francisco, L., & Alves, M. C. (2012). *Accounting Information and Performance Measurement in a Nonprofit Organization*. Paper presented at the Performance Measurement and Management Control: Global Issues, Bingley UK.
- Freeman, R. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.
- Freeman, R., Harrison, J. S., Wicks, A. C., Parmar, B. L., & De Colle, S. (2010). *Stakeholder theory: The state of the art*. Cambridge, U.K. ; New York: Cambridge University Press.
- Freeman, R. (1994). The politics of stakeholder theory. *Business Ethics Quarterly*, 4(4), 409–421.
- Froelich, K. A. (1999). Diversification of Revenue Strategies: Evolving Resource Dependence in Nonprofit Organizations. *Nonprofit & Voluntary Sector Quarterly*, 28(3), 246-268.
- Fry, R. E. (1995). Accountability in organizational life: Problem or opportunity for nonprofits? *Nonprofit Management and Leadership*, 6(2), 181-195.
- Gandolfi, F. (2005). How do organizations implement downsizing?—An Australian and New Zealand study. *Contemporary Management Research*, 1(1), 57-68.
- Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-GRAPH: Tutorial and annotated example. *Communications of the Association for Information Systems*, 16, 91-109.
- Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*, 4(1), 7.
- Geisser, S. (1974). A predictive approach to the random effect model. *Biometrika*, 61(1), 101-107.
- Gerber, S. B., & Finn, K. V. (2013). *Using SPSS for Windows: Data analysis and graphics*. New York: Springer.
- Gibbert, M., Ruigrok, W., & Wicki, B. (2008). What passes as a rigorous case study? *Strategic management journal*, 29(13), 1465-1474.
- Gómez, J. G., Costa, M. M., & Martínez-Lorente, Á. R. (2011). A critical evaluation of the EFQM model. *International Journal of Quality & Reliability Management*, 28(5), 484-502.
- Google. (2015, 10th November 2014). Google Forms. Retrieved 5 October, 2014, from <https://www.google.co.nz/forms/about/>
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse education today*, 24(2), 105-112.
- Grant, R. M. (1991). The resource-based theory of competitive advantage: Implications for strategy formulation. *California management review*, 33(3), 114-135.
- Gray, R., Owen, D., & Dams, C. (1996). *Accounting and Accountability: Changes and Challenges in Corporate Social and Environmental Reporting* Upper Saddle River: NJ: Prentice Hall.
- Greatbanks, R., & Tapp, D. (2007). The impact of balanced scorecards in a public sector environment: Empirical evidence from Dunedin City Council, New Zealand. *International Journal of Operations & Production Management*, 27(8), 846-873.
- Green, A., & Power, M. (2006). Defining transnationalism boundaries: New Zealand migrants in Australia. *Australian Journal of Communication*, 33(1), 35.
- Green, E. J., McCollum, V. C., & Hays, D. (2008). Teaching advocacy counseling within a social justice framework: Implications for school counselors and educators. *Journal for Social Action in Counseling and Psychology*, 1(2), 14-30.

- Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate behavioral research*, 26(3), 499-510.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational evaluation and policy analysis*, 11(3), 255-274.
- Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological methods*, 6(4), 430.
- Grigg, N., & Mann, R. (2008). Promoting excellence: An international study into creating awareness of business excellence models. *TQM Journal*, 20(3), 233-248. doi: 10.1108/17542730810867254
- Grigg, N. P., & Jayamaha, N. P. (2014). Measuring quality and excellence: On the role of structural equation modelling in quality and business excellence research. In P. K. Chopra (Ed.), *Quality, excellence and statistical measurement: A tribute to professor Gopal K. Kanji*. London: Wisdom House.
- Grigoroudis, E., Orfanoudaki, E., & Zopounidis, C. (2012). Strategic performance measurement in a healthcare organisation: A multiple criteria approach based on balanced scorecard. *Omega International Journal of Management Science*, 40(1), 104 - 119.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). Hoboken, New Jersey: John Wiley & Sons.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*, 29(2), 75-91.
- Guba, E. G. (1990). *The paradigm dialog*. USA: Sage.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.
- Gurd, B., & Gao, T. (2007). Lives in the balance: An analysis of the balanced scorecard (BSC) in healthcare organizations. *International Journal of Productivity and Performance Management*, 57(1), 6-21.
- Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. *Journal of Applied psychology*, 60(2), 159.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Calif.: Sage.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2012). Partial Least Squares: The Better Approach to Structural Equation Modeling? *Long Range Planning*, 45(5-6), 312-319. doi: 10.1016/j.lrp.2012.09.011
- Haley, M. (2014). Information technology and the quality improvement in defense industries. *TQM Journal*, 26(4), 348-359. doi: 10.1108/TQM-01-2014-0013
- Hall, P. D. (2005). Historical perspectives on nonprofit organizations in the United States. *The Jossey-Bass handbook of nonprofit leadership and management*, 2, 3-38.
- Hansmann, H. B. (1980). The Role of Nonprofit Enterprise. *The Yale Law Journal*, 89(5), 835-901. doi: 10.2307/796089
- Harley, K. (2005). Learning from Logframes: Reflections on Three Educational Development Projects in East and Southern Africa. *Compare A Journal of Comparative Education*, 35(1), 27-42.

- Hatch, M. J., & Cunliffe, A. L. (2013). *Organization theory : Modern, symbolic, and postmodern perspectives* (3 ed.). UK: Oxford University Press.
- Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment*, 7(3), 238.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods*, 7(2), 191-205.
- He, Z., Hill, J., Wang, P., & Yue, G. (2011). Validation of the theoretical model underlying the Baldrige criteria: Evidence from China. *Total Quality Management*, 22(2), 243-263.
- Heimovics, R. D., Herman, R. D., & Coughlin, C. L. J. (1993). Executive leadership and resource dependence in nonprofit organizations: A frame analysis. *Public Administration Review*, 53(5), 419-427.
- Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and evaluation in counseling and development*, 34(3), 177.
- Herman, R. D., & Renz, D. O. (2004). Doing things right: Effectiveness in local nonprofit organizations, a panel study. *Public Administration Review*, 694-704.
- Herzberg, F. (1965). The motivation to work among Finnish supervisors. *Personnel Psychology*, 18(4), 393-402.
- Hoerl, R., & Snee, R. D. (2012). *Statistical thinking: Improving business performance* (Vol. 48). Hoboken, NJ: John Wiley & Sons.
- Holste, J. S., & Fields, D. (2010). Trust and tacit knowledge sharing and use. *Journal of Knowledge Management*, 14(1), 128-140.
- Hoque, Z. (2014). 20 years of studies on the balanced scorecard: Trends, accomplishments, gaps and opportunities for future research. *The British Accounting Review*, 46(1), 33-59. doi: <http://dx.doi.org/10.1016/j.bar.2013.10.003>
- Hoque, Z., & Adams, C. (2011). The rise and use of Balanced Scorecard measures in Australian government departments. *Financial Accountability & Management*, 27(3), 308-334. doi: 10.1111/j.1468-0408.2011.00527.x
- Hoque, Z., & Parker, L. (2015). Accountability, governance and performance management in nonprofit organizations: An overview. In Z. Hoque & L. Parker (Eds.), *Performance Management in Nonprofit Organizations: Global Perspectives* (pp. 3-16). USA: Routledge.
- Hossain, M. M., & Prybutok, V. R. (2014). An empirical investigation of the Malcolm Baldrige National Quality Award framework using causal Latent Semantic Analysis. *International Journal of Business Excellence*, 7(2), 148-167.
- Hough, A. D., McGregor-Lowndes, M., & Ryan, C. (2015). Board Monitoring and the Balanced Scorecard in Nonprofits. In Z. Hoque & L. Parker (Eds.), *Performance Management in Nonprofit Organisations- Global Perspectives* (pp. 418). USA: Routledge.
- Houghton, C., Hunter, A., & Meskell, P. (2012). Linking aims, paradigm and method in nursing research. *Nurse Researcher*, 20(2), 34-39.
- Howe, K. R. (2012). Mixed Methods, Triangulation, and Causal Explanation. *Journal of Mixed Methods Research*, 6(2), 89-96. doi: 10.1177/1558689812437187
- Hoyle, R. H., Harris, M. J., & Judd, C. M. (2002). *Research methods in social relations* (7th ed.). USA: Thomson Learning.

- Hume, C., Clarke, P., & Hume, M. (2012). The role of knowledge management in the large non profit firm: Building a framework for KM success. *International Journal of Organisational Behaviour*, 17(3), 82-104.
- Iparraquirre, J. L., & Ma, R. (2015). Efficiency in the Provision of Social Care for Older People: A Three-Stage Data Envelopment Analysis Using Self-Reported Quality of Life. *Socio-Economic Planning Sciences*, 49, 33-46.
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of consumer research*, 30(2), 199-218.
- Jayamaha, N., Grigg, N., & Mann, R. (2009). A study of the validity of three major business excellence models in the Asia Pacific region. *Total Quality Management & Business Excellence*, 20(11), 1213-1227. doi: 10.1080/14783360903247536
- Jayamaha, N. P., & Grigg, N. P. (2014). Measuring quality and excellence: On the role of structural equation modelling in quality and business excellence research In P. Chopra (Ed.), *Quality, Excellence and Measurement: A Tribute to Professor Gopal K. Kanji* (pp. 323-348). Leeds, UK: Wisdom House.
- Jayamaha, N. P., Grigg, N. P., & Mann, R. S. (2011). Empirical analysis of the Baldrige Criteria as both an organisational performance measure and a theoretical model. *Measuring Business Excellence*, 15(1), 20-33.
- Jehn, K. A., & Jonsen, K. (2010). A multimethod approach to the study of sensitive organizational issues. *Journal of Mixed Methods Research*, 4(4), 313-341. doi: 10.1177/1558689810380920
- Jensen, M. C. (2002). Value Maximization, Stakeholder Theory, and the Corporate Objective Function. *Business Ethics Quarterly*, 12(2), 235-256.
- Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative science quarterly*, 24(4), 602-611.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), 14-26.
- Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika*, 57(2), 239-251.
- Kanter, R. M., & Summers, D. V. (1994). *Doing well while doing good: Dilemmas of performance measurement in nonprofit organizations and the need for a multiple-constituency approach*. Londres: Sage.
- Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 353-370.
- Kaplan, R. S. (2002). *The balanced scorecard and nonprofit organizations*. USA: Harvard Business School Publishing.
- Kaplan, R. S. (2008). Conceptual Foundations of the Balanced Scorecard. In S. C. Christopher, A. G. Hopwood, & M. D. Shields (Eds.), *Handbooks of Management Accounting Research* (Vol. 3, pp. 1253-1269): Elsevier.
- Kaplan, R. S. (2012). The balanced scorecard: Comments on balanced scorecard commentaries. *Journal of Accounting and Organizational Change*, 8(4), 539-545.
- Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard—measures that drive performance. *Harvard business review*, 70(1), 71-79.
- Kaplan, R. S., & Norton, D. P. (1996a). Linking the Balanced Scorecard to strategy. *California management review*, 39(1), 53-79.
- Kaplan, R. S., & Norton, D. P. (1996b). Using the balanced scorecard as a strategic management system. On *Harvard business review*.

- Kaplan, R. S., & Norton, D. P. (2001). Transforming the Balanced Scorecard from performance measurement to strategic management: Part I. *Accounting Horizons*, 15(1), 87-104. doi: 10.2308/acch.2001.15.1.87
- Kaplan, R. S., & Norton, D. P. (2004). The strategy map: Guide to aligning intangible assets. *Strategy & Leadership*, 32(5), 10-17.
- Kast, F. E., & Rosenzweig, J. E. (1972). General System Theory: Applications for Organization and Management. *Academy of Management Journal*, 15(4), 447-465. doi: 10.2307/255141
- Kearns, K. P. (1996). *Managing for accountability: Preserving the public trust in public and nonprofit organizations*: Jossey-Bass.
- Kellogg Foundation. (2006). *Logic model development guide*.
- Kendall, J., & Knapp, M. (2000). Measuring the performance of voluntary organizations. *Public Management Review*, 2(1), 105-132.
- Kennerley, M., & Neely, A. (2002). A framework of the factors affecting the evolution of performance measurement systems. *International Journal of Operations & Production Management*, 22(11), 1222-1245. doi: 10.1108/01443570210450293
- Kilby, P. (2006). Accountability for empowerment: Dilemmas facing non-governmental organizations. *World development*, 34(6), 951-963.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3 ed.). NY: Guilford Press.
- Knox, S., & Gruar, C. (2007). The application of stakeholder theory to relationship marketing strategy development in a non-profit organization. *Journal of Business Ethics*, 75(2), 115-135.
- Kocakülâh, M. C., & Austill, A. D. (2007). Balanced Scorecard Application in the Health Care Industry: A Case Study. *Journal of Health Care Finance*, 34(1), 72-99.
- Kong, E. (2008). The development of strategic management in the non-profit context: Intellectual capital in social service non-profit organizations. *International Journal of Management Reviews*, 10(3), 281-299. doi: 10.1111/j.1468-2370.2007.00224.x
- Kong, E. (2010). Analyzing BSC and IC's usefulness in nonprofit organizations. *Journal of intellectual capital*, 11(3), 284-304.
- Koufteros, X., Vergheze, A., & Lucianetti, L. (2014). The effect of performance measurement systems on firm performance: A cross-sectional and a longitudinal study. *Journal of Operations Management*, 32, 313-336. doi: 10.1016/j.jom.2014.06.003
- Krashinsky, M. (1997). Stakeholder theories of the non-profit sector: One cut at the economic literature. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 8(2), 149-161.
- Laratta, R. (2010). Accountability and Advocacy in Nonprofit Organisations: A Case Study. *The Nonprofit Review*, 10(1), 67-79.
- Lawrie, G., & Cobbold, I. (2004). Third-generation balanced scorecard: Evolution of an effective strategic control tool. *International Journal of Productivity and Performance Management*, 53(7), 611-623. doi: 10.1108/17410400410561231
- Lawry, R. P. (1995). Accountability and nonprofit organizations: An ethical perspective. *Nonprofit Management and Leadership*, 6(2), 171-180.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.

- LeRoux, K. (2009). Managing stakeholder demands: Balancing responsiveness to clients and funding agents in nonprofit social service organizations. *Administration & Society*, 41(2), 158-184. doi: 10.1177/0095399709332298
- LeRoux, K., & Wright, N. S. (2010). Does performance measurement improve strategic decision making? Findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 571-587.
- Levitt, B., & March, J. G. (1988). Organizational learning. *Annual review of sociology*, 319-340.
- Limburg, D., Knowles, C., & McCulloch, M. (2012, 26-28th March). *Meeting the information needs of charity trustees: Can Enterprise Performance Management Systems help?* Paper presented at the UK Academy for Information Systems Oxford University, UK.
- Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretive research. *Qualitative inquiry*, 1(3), 275-289.
- Lipsey, M. W. (2000). Statistical Conclusion Validity. In L. Bickman (Ed.), *Validity and social experimentation: Donald Campbell's legacy* (1 ed., pp. 101). USA: Sage.
- Love, A. (2004). Implementation Evaluation. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed.). San Francisco: Jossey-Bass.
- Lynch-Cerullo, K., & Cooney, K. (2011). Moving from outputs to outcomes: A review of the evolution of performance measurement in the Human service nonprofit sector. *Administration in Social Work*, 35(4), 364-388.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research*, 35(6), 382-386.
- Lyons, M., Hocking, S., Hems, L., & Salamon, L. M. (1999). Australia Global civil society. In L. M. Salamon, H. K. Anheier, R. List, S. Toepler, & S. W. Sokolowski (Eds.), *Global civil society: Dimensions of the Nonprofit sector* (pp. 203-217). Baltimore, USA.
- Maas, K., & Liket, K. (2011). Social Impact Measurement: Classification of Methods. In R. L. Burritt, S. Schaltegger, M. Bennett, T. Pohjola, & M. Csutora (Eds.), *Environmental Management Accounting and Supply Chain Management* (Vol. 27, pp. 171-202). New York: Springer.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2), 130-149. doi: 10.1037/1082-989x.1.2.130
- MacIndoe, H., & Barman, E. (2013). How organizational stakeholders shape performance measurement in nonprofits: Exploring a multidimensional measure. *Nonprofit and Voluntary Sector Quarterly*, 42(4), 716-738.
- Madden, K., & Scaife, W. (2010). Philanthropy in Australia. *International Encyclopaedia of Civil Society* (pp. 1192-1196): Springer.
- Mak, L., & Marshall, S. K. (2004). Perceived mattering in young adults' romantic relationships. *Journal of Social and Personal Relationships*, 21(4), 469-486.
- Malatesta, D., & Smith, C. R. (2014). Lessons from resource dependence theory for contemporary public and nonprofit management. *Public Administration Review*, 74(1), 14-25.
- Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: From constructs to theory. *Journal of Operations Management*, 16(4), 407-425.

- Maltz, A. C., Shenhar, A. J., & Reilly, R. R. (2003). Beyond the balanced scorecard: Refining the search for organizational success measures. *Long Range Planning*, 36(2), 187-204.
- Martello, M., Watson, J. G., & Fischer, M. J. (2011). Implementing a Balanced Scorecard in a not-for-profit organization. *Journal of Business & Economics Research*, 6(9), 67-80.
- May, D. R., Gilson, R. L., & Harter, L. M. (2004). The psychological conditions of meaningfulness, safety and availability and the engagement of the human spirit at work. *Journal of occupational and organizational psychology*, 77(1), 11-37.
- McDermott, R., & O'Dell, C. (2001). Overcoming cultural barriers to sharing knowledge. *Journal of Knowledge Management*, 5(1), 76-85. doi: 10.1108/13673270110384428
- McDonald, R. E. (2007). An investigation of innovation in nonprofit organizations: The role of organizational mission. *Nonprofit and Voluntary Sector Quarterly*, 36(2), 256-281.
- McDonald, R. P. (2013). *Test theory: A unified treatment*. New York: Routledge.
- McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review*, 26(1), 117-127.
- Mellat-Parast, M. (2015). A longitudinal assessment of the linkages among the Baldrige criteria using independent reviewers' scores. *International Journal of Production Economics*, 164, 24-34.
- Mendibil, K., & MacBryde, J. (2005). Designing effective team-based performance measurement systems: An integrated approach. *Production Planning & Control*, 16(2), 208-225. doi: 10.1080/09537280512331333101
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), 741.
- Meyer, M. W. (2002). *Rethinking performance measurement : Beyond the balanced scorecard*. New York: Cambridge University Press.
- Milne, P. (2007). Motivation, incentives and organisational culture. *Journal of Knowledge Management*, 11(6), 28-38. doi: 10.1108/13673270710832145
- Mitchell, T. R. (1985). An evaluation of the validity of correlational research conducted in organizations. *Academy of Management Review*, 10(2), 192-205.
- Moore, M. H. (2003). The Public Value Scorecard: A Rejoinder and an Alternative to 'Strategic Performance Measurement and Management in Non-Profit Organizations' by Robert Kaplan (May 2003). *Hauser Center for Nonprofit Organizations Working Paper*, 18.
- Morgan, G. G. (2010). The use of charitable status as a basis for regulation of nonprofit accounting. *Voluntary Sector Review*, 1(2), 209-232.
- Morse, J. M. (1998). Validity by committee. *Qualitative health research*, 8(4), 443-445.
- Mount, J. (1996). Why donors give. *Nonprofit Management and Leadership*, 7(1), 3-14.
- Moxham, C. (2009). Performance measurement: Examining the applicability of the existing body of knowledge to nonprofit organisations. *International Journal of Operations & Production Management*, 29(7), 740-763.
- Mulgan, G. (2010). Measuring social value. *Stanford Soc Innov Rev*, 8(3), 38-43.
- Myers, K. K., & Oetzel, J. G. (2003). Exploring the dimensions of organizational assimilation: Creating and validating a measure. *Communication Quarterly*, 51(4), 438-457. doi: 10.1080/01463370309370166
- Myers, M. D. (2013). *Qualitative research in business and management* (2nd ed.). London: Sage.

- Najam, A. (2000). The Four C's of Government Third Sector-Government Relations. *Nonprofit Management and Leadership*, 10(4), 375-396.
- Najmi, M., Etebari, M., & Emami, S. (2012). A framework to review Performance Prism. *International Journal of Operations & Production Management*, 32(10), 1124-1146. doi: 10.1108/01443571211274486
- Nardi, P. M. (2015). *Doing survey research* (3rd ed.). New York, USA: Routledge.
- Neely, A., Adams, C., & Crowe, P. (2001). The performance prism in practice. *Measuring Business Excellence*, 5(2), 6-13.
- Neely, A., Gregory, M., & Platts, K. (2005). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management*, 25(12), 1128-1263. doi: 10.1108/01443570510633639
- Neely, A., Marr, B., Roos, G., Pike, S., & Gupta, O. (2003). Towards the third generation of performance measurement. *Controlling*, 15(3/4), 129-135.
- Neely, A. D. (2007). Measuring performance: The operations management perspective. In A. D. Neely (Ed.), *Business Performance Measurement : Unifying Theories and Integrating Practice* (2nd ed., pp. 64-81). Cambridge ; New York: Cambridge University Press.
- Neely, A. D., Adams, C., & Kennerley, M. (2002). *The performance prism : The Scorecard for Measuring and Managing Business Success*. London: Prentice Hall.
- Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, 22(4), 287-293.
- National Institute of Standards and Technology, (NIST), (2015). Baldrige excellence framework: 2015-2016. Gaithersburg, MD: National Institute of Standards and Technology.
- Niven, P. R. (2008). *Balanced Scorecard Step-by-step for Government and Nonprofit Agencies*. Hoboken, N.J.: J. Wiley & Sons.
- Noell, C., & Lund, M. (2002). *The Balanced Scorecard (BSC) for Danish Farms—Vague Framework or Functional Instrument?* Paper presented at the Farm Management. Proceedings of NJF Seminar No. 345.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in health sciences education*, 15(5), 625-632.
- Nørreklit, H. (2000). The balance on the balanced scorecard a critical analysis of some of its assumptions. *Management Accounting Research*, 11(1), 65-88. doi: 10.1006/mare.1999.0121
- Nørreklit, H., Nørreklit, L., Mitchell, F., & Bjørnenak, T. (2012). The rise of the balanced scorecard! Relevance regained? *Journal of Accounting & Organizational Change*, 8(4), 490-510. doi: 10.1108/18325911211273491
- Northcott, D., & Taulapapa, T. M. a. (2012). Using the balanced scorecard to manage performance in public sector organizations: Issues and challenges. *International Journal of Public Sector Management*, 25(3), 166 - 191. doi: 10.1108/09513551211224234
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3 ed.). New York: McGraw-Hill.
- O'Cathain, A. (2010). Assessing the Quality of Mixed Methods Research: Towards a Comprehensive Framework. In A. Tashakkori & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2 ed., pp. 531-580). LA: Sage.
- O'Neill, M., & Young, D. R. (1988). *Educating managers of nonprofit organizations*. New York: Praeger.

- Oakland, J. S. (2001). *Total organizational excellence: Achieving world-class performance*. Woburn, MA: Butterworth-Heinemann.
- Oakland, J. S. (2004). *Oakland on quality management*. Boston: Elsevier/Butterworth-Heinemann.
- Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed Research as a Tool for Developing Quantitative Instruments. *Journal of Mixed Methods Research, 4*(1), 56-78. doi: 10.1177/1558689809355805
- Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools, 13*(1), 48-63.
- Orlikowski, W. J., & Baroudi, J. J. (2002). Studying information technology in organizations : Research approaches and assumptions. In M. D. Myers & D. Avison (Eds.), *Qualitative research in information systems: A Reader* (pp. 51-77). London: Sage.
- Ospina, S., Diaz, W., & O'Sullivan, J. F. (2002). Negotiating accountability: Managerial lessons from identity-based nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly, 31*(1), 5-31.
- Ostrower, F. (2007). Nonprofit governance in the United States: Findings on performance and accountability from the first national representative study.
- Otley, D. (1999). Performance management: A framework for management control systems research. *Management Accounting Research, 10*(4), 363-382.
- Otley, D. (2001). Accounting performance measurement : A review of its purposes and practices. *International Journal of Business Performance Management, 3*, 245-260.
- Pannirselvam, G. P., Siferd, S. P., & Ruch, W. A. (1998). Validation of the Arizona Governor's Quality Award criteria: A test of the Baldrige criteria. *Journal of Operations Management, 16*(5), 529-550.
- Parmenter, D. (2012). *A table without any legs: A critique of the balanced scorecard methodology*. In implementing Winning KPIs whitepaper. Retrieved from <http://davidparmenter.com/how-toguides>
- Parmenter, D. (2015). *Key performance indicators: Developing, Implementing, and Using Winning KPIs*. Hoboken, New Jersey: John Wiley & Sons.
- Paton, R. (2003). *Managing and measuring social enterprises*. London: Sage.
- Peng, A. C., Lin, H.-E., Schaubroeck, J., McDonough, E. F., Hu, B., & Zhang, A. (2016). CEO Intellectual Stimulation and Employee Work Meaningfulness The Moderating Role of Organizational Context. *Group & Organization Management, 41*(2), 203-231.
- Perkins, M., Grey, A., & Remmers, H. (2014). What do we really mean by 'Balanced Scorecard'? *International Journal of Productivity and Performance Management, 63*(2), 148-169. doi: 10.1108/IJPPM-11-2012-0127
- Pessanha, D. S. d. S., & Prochnik, V. (2006). Practitioners' Opinions on Academics' Critics on the Balanced Scorecard. Retrieved 28.01.2016, from <http://ssrn.com/abstract=1094308>
- Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *Mis Quarterly, 31*(4), 623-656.
- Podsakoff, M. P., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology, 63*, 539-569.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied psychology, 88*, 879-903.

- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of management*, 12(4), 531-544.
- Poister, T. H. (2003). *Measuring performance in public and nonprofit organizations*. San Francisco, Calif.
- Poole, D. L., Nelson, J., Carnahan, S., Chepenik, N. G., & Tubiak, C. (2000). Evaluating performance measurement systems in nonprofit agencies: The program accountability quality scale (PAQS). *American Journal of Evaluation*, 21(1), 15-26.
- Popper, K. R. S. (2002). *The logic of scientific discovery*. London: Routledge.
- Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard Business Review*, 86(1), 78-93.
- Pruijt, H. (2012). InterviewStreamliner, a minimalist, free, open Source, relational approach to computer-assisted qualitative data analysis software. *Social Science Computer Review*, 30(2), 248-253.
- Prybutok, V. R., Zhang, X., & Ryan, S. D. (2008). Evaluating leadership, IT quality, and net benefits in an e-government environment. *Information & Management*, 45(3), 143-152.
- Qu, S. Q., & Cooper, D. J. (2011). The role of inscriptions in producing a balanced scorecard. *Accounting, Organizations and Society*, 36(6), 344-362.
- Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, N.J. : Lawrence Erlbaum Associates.
- Renz, D. O. (2007). Nonprofit Governance and the Work of the Board. *Midwest Center for Nonprofit Leadership*.
- Renz, D. O., & Anderson, F. O. (2014). Nonprofit governance: A review of the field. In C. Cornforth & W. A. Brown (Eds.), *Nonprofit governance: Innovative perspectives and approaches* (pp. 17-46). Abingdon, Oxon: Routledge.
- Richmond, B. J., Mook, L., & Jack, Q. (2003). Social accounting for nonprofits: Two models. *Nonprofit Management and Leadership*, 13(4), 308-324.
- Ridwan, R., Harun, H. Y., & Fahmid, I. M. (2013). The Impact of the Balanced Scorecard on Corporate Performance: The Case of an Australian Public Sector Enterprise. *International Business Research*, 6(10), 103-110.
- Riege, A. M. (2003). Validity and reliability tests in case study research: A literature review with "hands-on" applications for each research phase. *An International Journal of Qualitative Market Research*, 6(2), 75-86.
- Rigdon, E. E., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. *Review of marketing research*, 7(7), 255-296.
- Rillo, M. (2004, 30-31 January). *Limitations of balanced scorecard*. Paper presented at the 2nd International Conference on Business Administration, Parnu.
- Ritchie, W. J., & Kolodinsky, R. W. (2003). Nonprofit organization financial performance measurement: An evaluation of new and existing financial performance measures. *Nonprofit Management and Leadership*, 13(4), 367-381.
- Roberts, N., & Thatcher, J. (2009). Conceptualizing and testing formative constructs: Tutorial and annotated example. *ACM SigMIS Database*, 40(3), 9-39. doi: 10.1145/1592401.1592405
- Rorty, R. (1991). *Essays on Heidegger and others: Philosophical papers* (Vol. 2). New York, USA: Cambridge University Press.
- Rungtusanatham, M. J., Choi, T. Y., Hollingworth, D. G., Wu, Z., & Forza, C. (2003). Survey research in operations management: Historical analyses. *Journal of Operations Management*, 21(4), 475-488.

- Salamon, L. M. (1987). Of market failure, voluntary failure, and third-party government: Toward a theory of government-nonprofit relations in the modern welfare state. *Nonprofit and Voluntary Sector Quarterly*, 16(1-2), 29-49.
- Salamon, L. M. (1994). The rise of the nonprofit sector. *Foreign Affairs*, 73, 109-122.
- Salamon, L. M., & Anheier, H. K. (1997). *Defining the nonprofit sector : A Cross-National Analysis*. New York: Manchester University Press.
- Salamon, L. M., & Sokolowski, S. W. (2004). *Global civil society: Dimensions of the nonprofit sector* (Vol. 2). Bloomfield: Kumarian
- Salamon, L. M., Sokolowski, S. W., Haddock, M. A., & Tice, H. S. (2013). The state of global civil society and volunteering: Latest findings from the implementation of the UN nonprofit handbook. *Center for Civil Society Studies Working Paper*(49).
- Salancik, G. R., & Pfeffer, J. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Los Angeles: Sage.
- Salterio, S. (2012). Balancing the scorecard through academic accounting research: opportunity lost? *Journal of Accounting & Organizational Change*, 8(4), 458-474.
- Samuelsson, P., & Nilsson, L.-E. (2002). Self-assessment practices in large organisations: Experiences from using the EFQM excellence model. *International Journal of Quality & Reliability Management*, 19(1), 10-23.
- Sanders, J., O'Brien, M., Tennant, M., Sokolowski, S. W., & Salamon, L. M. (2008). *The New Zealand non-profit sector in comparative perspective*. Wellington: Office for the Community and Voluntary Sector
- Saunders, M., & Mann, R. (2005). Self-assessment in a multi-organisational network. *International Journal of Quality & Reliability Management*, 22(6), 554-571.
- Saunders, M., Mann, R., & Grigg, N. (2008). Review processes for improving business excellence frameworks. *International Journal of Quality & Reliability Management*, 25(9), 928-942.
- Saunders, M., Mann, R., & Smith, R. (2007). Benchmarking strategy deployment practices. *Benchmarking: An International Journal*, 14(5), 609-623.
- Saunders, M. N. K. (2012). Choosing research participants. In G. Symon & C. Cassell (Eds.), *Qualitative organizational research : Core methods and current challenges* (pp. 37-55). London: Sage.
- Saunders, M. N. K., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6 ed.). NY: Pearson.
- Sawhill, J. C., & Williamson, D. (2001). Mission impossible?: Measuring success in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 371-386.
- Scaife, W., McDonald, K., Williamson, A., & Mossel, V. (2015). Giving in Australia: Philanthropic Potential Beginning to Be Realized. In P. Wierking & F. Handy (Eds.), *The Palgrave Handbook of Global Philanthropy* (pp. 488-505). UK: Palgrave Macmillan.
- Schermerhorn, J. R. (2011). *Management* (11 ed.). Hoboken, NJ: John Wiley & Sons.
- Schläfke, M., Silvi, R., & Möller, K. (2013). A framework for business analytics in performance management. *International Journal of Productivity and Performance Management*, 62(1), 110-122. doi: 10.1108/17410401311285327
- Schneider, J. A. (2009). Organizational social capital and nonprofits. *Nonprofit and Voluntary Sector Quarterly*, 38(4), 643-662.

- Shekar, A. (2007). An Innovative Model of Service Development: A process guide for service managers. *The Innovation Journal: The Public Sector Innovation Journal*, 12(1), 1-18.
- Simon, H. A. (1957). *Models of man: Social and rational; mathematical essays on rational human behavior in society setting*. New York, NY: Wiley.
- Smith, B. R., Cronley, M. L., & Barr, T. F. (2012). Funding implications of social enterprise: The role of mission consistency, entrepreneurial competence, and attitude toward social enterprise on donor behavior. *Journal of Public Policy & Marketing*, 31(1), 142-157.
- Smulowitz, S. (2015). Evidence for the performance prism in higher education. *Measuring Business Excellence*, 19(1), 70-80. doi: 10.1108/MBE-11-2014-0043
- Solomon, J. (2010). *Corporate governance and accountability* (3rd ed.). Chichester, West Sussex, U.K.: Wiley.
- Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2013a, 30 October-1 November). *Examining the performance measurement in the Australasian nonprofit sector*. Paper presented at the Performance Management Association of Australasia (PMAA) Conference 2013, Queenstown, New Zealand.
- Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2013b, 20-21 June). *Performance measurement in the nonprofit sector*. Paper presented at the 11th ANZAM Operations, Supply Chain and Services Management Symposium, Research Industry Partnership Brisbane.
- Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2014, 3 - 4 July). *Developing a conceptual model for Measuring Performance of Australasian Nonprofit Sector*. Paper presented at the 12th ANZAM Operations, Supply Chain and Services Management Symposium, University of Auckland Business School.
- Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2016a, 26-29 June). *Developing a Performance Scoring System for Nonprofit Organisations*. Paper presented at the 10th conference of the Performance Measurement Association, Edinburgh, Scotland.
- Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2016b). Operationalising Performance Dimensions for the Australasian Nonprofit Healthcare Sector. *TQM Journal*, 28(6), forthcoming.
- Speckbacher, G. (2008). Nonprofit versus corporate governance: An economic approach. *Nonprofit Management and Leadership*, 18(3), 295-320.
- Srikanthan, G., & Dalrymple, J. (2003). Developing alternative perspectives for quality in higher education. *International Journal of Educational Management*, 17(3), 126-136.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks: Sage
- Statistics New Zealand. (2016). Retrieved January 10, 201, from <http://www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/non-profit-organisation.aspx>
- Steinberg, R. (2006). Economic theories of nonprofit organisations. In R. Steinberg & W. W. Powell (Eds.), *The Nonprofit Sector : A Research Handbook* (2nd ed., pp. 117-139). New Haven, CT: Yale University Press.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science, New Series*, 103(2684), 677-680.
- Stewart, L. J., & Bestor, W. E. (2000). Applying a Balanced Scorecard to Health Care Organizations. *Journal of Corporate Accounting & Finance* 11(3), 75-82.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society*, 36(2), 111-147.

- Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*(24), 380-427.
- Straub, D. W. (1989). Validating instruments in MIS research. *Mis Quarterly, 13*(2), 147-169.
- Striteska, M., & Spickova, M. (2012). Review and comparison of performance measurement systems. *Journal of Organizational Management Studies, 2012*, 1-14. doi: 10.5171/2012.114900
- Talbot, C. (2010). *Theories of performance: Organizational and service improvement in the public domain*. Oxford, NY: Oxford University Press.
- Tangen, S. (2004). Performance measurement: From philosophy to practice. *International Journal of Productivity and Performance Management, 53*(8), 726-737.
- Tashakkori, A., & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research* (2 ed.). LA: Sage.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. USA: Sage.
- Tenenhaus, M., Amato, S., & Esposito Vinzi, V. (2004). *A global goodness-of-fit index for PLS structural equation modelling*. Paper presented at the Proceedings of the XLII SIS scientific meeting.
- Tennant, M., Castle, C., O'Brien, M., Salamon, L. M., & Sanders, J. (2006). *Defining the Nonprofit Sector: New Zealand*. Johns Hopkins University Center for Civil Society Studies.
- Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 3-23). USA: Sage.
- Tremblay-Boire, J., & Prakash, A. (2015). Accountability.org: Online Disclosures by U.S. Nonprofits. *International Journal of Voluntary and Nonprofit Organizations, 26*(2), 693-719. doi: 10.1007/s11266-014-9452-3
- Trevor-Roberts, E., Ashkanasy, N. M., & Kennedy, J. C. (2003). The egalitarian leader: A comparison of leadership in Australia and New Zealand. *Asia Pacific Journal of Management, 20*(4), 517-540.
- Trotta, A., Cardamone, E., Cavallaro, G., & Mauro, M. (2013). Applying the Balanced Scorecard approach in teaching hospitals: A literature review and conceptual framework. *The International Journal of Health Planning and Management, 28*(2), 181-201. doi: 10.1002/hpm.2132
- Tung, A., Baird, K., & Schoch, H. P. (2011). Factors influencing the effectiveness of performance measurement systems. *International Journal of Operations & Production Management, 31*(12), 1287-1310. doi: 10.1108/014435711111187457
- Valderrama, T. G., Cornejo, V. R., & Bordoy, D. R. (2013). Balanced Scorecard and Efficiency: Design and Empirical Validation of a Strategic Map in the University by Means of DEA. *American Journal of Operations Research, 3*(1), 30-52.
- Varadarajan, P. R., & Menon, A. (1988). Cause-related marketing: A coalignment of marketing strategy and corporate philanthropy. *The Journal of Marketing, 58*-74.
- Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research, 25*(1), 1-28.

- Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47(1), 65-72.
- Viader, A., & Espina, M. (2014). Are not-for-profits learning from for-profit-organizations? A look into governance. *Corporate Governance*, 14(1), 1-14.
- Vinson, E., & Morley, E. (2000). *Collection and use of outcome information by nonprofit organizations: Findings and recommendations*. Paper presented at the 29th Annual Conference of the Association for Research on Nonprofit Organizations and Voluntary Action, November, Washington, DC, The Urban Institute.
- Voss, C., Tsiriktsis, N., & Frohlich, M. (2002). Case research in operations management. *International Journal of Operations & Production Management*, 22(2), 195-219.
- Wahyuni, D. (2012). The Research Design Maze: Understanding Paradigms, Cases, Methods and Methodologies. *Journal of Applied Management Accounting Research*, 10(1), 69-80.
- Walsham, G. (1995). Interpretive case studies in IS research: Nature and method. *European Journal of information systems*, 4(2), 74-81.
- Watson, G. H. (1993). *Strategic benchmarking : How to rate your company's performance against the world's best*. New York: J.Wiley and Sons.
- Weisbrod, B. A. (1975). *Toward a theory of the voluntary non-profit sector in a three-sector economy*. Madison: Institute for Research on Poverty, University of Wisconsin.
- Weisbrod, B. A., & Dominguez, N. D. (1986). Demand for collective goods in private nonprofit markets: Can fundraising expenditures help overcome free-rider behavior? *Journal of public economics*, 30(1), 83-96.
- Welch, D. (2007). Counting Non-profit Institutions in New Zealand. *Sector Highlights* (pp. 16). Wellington: Statistics New Zealand.
- Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational & Psychological Measurement*, 34, 25-33.
- Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, 33(1), 177-195.
- Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490-495.
- Wholey, J. S. (1999). Performance-Based Management: Responding to the Challenges. *Public Productivity & Management Review*, 22(3), 288-307. doi: 10.2307/3380705
- Williams, D. W. (2003). Measuring Government in the Early Twentieth Century. *American Society for Public Administration Review*, 63(6), 643-659. doi: 10.1111/1540-6210.00329
- Williams, D. W. (2004). Evolution of Performance Measurement Until 1930. *Administration & Society*, 36(2), 131-165. doi: 10.1177/0095399704263473
- Williams, W., & Lewis, D. (2008). Strategic management tools and public sector management: The challenge of context specificity. *Public Management Review*, 10(5), 653-671.
- Wilson, D. D., & Collier, D. A. (2000). An Empirical Investigation of the Malcolm Balridge National Quality Award Causal Model. *Decision Sciences*, 31(2), 361-390.

- Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce *Evaluation of econometric models* (pp. 47-74): Academic Press.
- Wold, H. (1982). Soft modelling: The basic design and some extensions. In H. Wold & K. G. Joreskog (Eds.), *Systems under indirect observation*, (Vol. 2, pp. 1-54). North holland: Amsterdam.
- Woods, S. A., & Sofat, J. A. (2013). Personality and engagement at work: The mediating role of psychological meaningfulness. *Journal of applied social psychology*, 43(11), 2203-2210.
- Wu, S., & Hung, J.-M. (2008). A performance evaluation model of CRM on non-profit organisations. *Total Quality Management & Business Excellence*, 19(4), 321-342. doi: 10.1080/14783360701591978
- Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, 25(5), 508-518.
- Yadav, N., & Sagar, M. (2013). Performance measurement and management frameworks: Research trends of the last two decades. *Business Process Management Journal*, 19(6), 947-971.
- Yang, C. C., Cheng, L. Y., & Yang, C. W. (2005). A study of implementing Balanced Scorecard (BSC) in non-profit organizations: A case study of private hospital. *Human Systems Management*, 24(4), 285-300.
- Yin, R. K. (2014). *Case study research : Design and Methods* (5 ed.). LA: Sage.
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79-94.
- Youndt, M. A., Subramaniam, M., & Snell, S. A. (2004). Intellectual capital profiles: An examination of investments and returns. *Journal of Management studies*, 41(2), 335-361.
- Young, D. R., Bania, N., & Bailey, D. (1996). Structure and accountability: A study of national nonprofit associations. *Nonprofit Management and Leadership*, 6(4), 347-365. doi: 10.1002/nml.4130060405
- Zikmund, W. G. (2013). *Business research methods* (9th international ed.). South-Western, Australia: Cengage Learning.
- Zuckerman, A. M. (2012). *Healthcare strategic planning* (3rd ed.). Chicago: Health Administration Press.

APPENDICES

- APPENDIX A: The Semi-Structured Questionnaire Used in the Case Study
- APPENDIX B: Invitation (Email) to Participants: the Quantitative Questionnaire
- APPENDIX C: The Online Survey Questionnaire
- APPENDIX D: Massey University Human Ethics Approval – Low Risk
- APPENDIX E: Quality Assessment of the Ancillary Model
- APPENDIX F: The Author’s Publications Related to her Doctoral Research: the DRC 16 Forms
- APPENDIX G: Researcher’s Journal Articles (Page 1 only)

APPENDIX A: THE SEMI-STRUCTURED QUESTIONNAIRE USED IN THE CASE STUDY



MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

.....
.....
.....

Dear

Performance Measurement in Australasian Nonprofit Healthcare Organisations: An Exploratory Survey

I'm writing to request your valuable input into my PhD research project. I am Ishani Soysa, a doctoral student (PhD) from Massey University, Palmerston North. My supervisors are Dr. Nihal Jayamaha and Associate Prof. Nigel Grigg. My aim is to develop a performance measurement framework that can be easily customised to suit individual non-profit healthcare organisations, given their missions, strategic goals and objectives.

This questionnaire is designed as a part of my research work, mainly aiming to gain insights about the stakeholder environment, strategic goals and objectives and the organisational processes that are used by community healthcare nonprofit organisations to deliver value to the stakeholders. Through this I hope to develop a model that will be subsequently tested using survey data collected from Australian and New Zealand nonprofit organisations (community healthcare).

All the data collected from this questionnaire will be used for my PhD research only and the anonymity of the respondents and their organisations will be preserved at all times. Please look through the questions carefully, and provide your responses as descriptively as you can, where necessary. Do please use the additional blank pages (pages 7 and 8) that I have provided for your comments. I would greatly appreciate your participation in my study. In return for the valuable data that you provide me by way of responding to this questionnaire, I will send you a copy of my final complete performance measurement instrument and an executive summary of my interim findings.

This research is conducted in accordance with Massey University Human Ethics Committee guidelines on low risk research. If you have any queries on completing the questionnaire do not hesitate to contact me. I look forward to hearing from soon.

Thanking you,
Yours faithfully,

Soysa H. S. B.

Ishani Soysa

Email: I.Soyasa@Massey.ac.nz Tel: (06)3569099, Ext 85696

Te Kunenga
ki Pārehuroa

School of Engineering and Advanced Technology Private Bag 11-222, Palmerston North 4442, New Zealand T +64 6 350 5525
F +64 6 350 5604 www.me.ac.nz

Performance Measurement in Australasian Nonprofit Healthcare Organisations: An Exploratory Survey

Instruction for completing the survey and Key definitions

- Answer each question by filling in the space provided.
- Please type as many words as you need to (from question 4 onwards) in the space provided; the sentences scroll upwards once the space is being populated.
- Please use the space provided in pages 8 and 9 for any additional comments that you would like to make.

Key definitions

Mission : A mission is the purpose of a company or organization, its reason for existence. The mission statement should guide the actions of the organisation, spell out its overall goal, provide a path, and guide decision-making. It provides "the framework or context within which the company's strategies are formulated." It's like a goal for what the company wants to do for the world.

Vision : An aspirational description of what an organization would like to achieve or accomplish in the mid-term or long-term future. It is intended to serve as a clear guide for choosing current and future courses of action.

Stakeholders : A group, organisation, member or system that affects or can be affected by an organisation's actions. E.g. Customers, Employees, Donors, Legislative bodies, Community, etc.

Processes : An activity or set of activities that will accomplish a specific organisational goal.

Performance Measurement: The process of defining, monitoring, and using objective indicators of performance of the organisation and its programmes on a regular basis. It involves creating a simple, but effective and efficient system for determining whether the organisation meet its objectives.

Performance Measurement in Australasian Nonprofit Healthcare Organisations: An Exploratory Survey

Part One: Basic Information and Background about You and Your Organisation

1. To which category mentioned below does your organisation belong to? Select only one category.

- | | | | |
|---------------------------------------|-----------------------|--------------------------------|-----------------------|
| Hospitals and rehabilitation | <input type="radio"/> | Nursing homes | <input type="radio"/> |
| Mental health and crisis intervention | <input type="radio"/> | Support and ancillary services | <input type="radio"/> |
| Other health services | <input type="radio"/> | Please specify | |

2. How long has your organisation (including any predecessor/s if applicable) been in operation?

- Less than 5 years
- 5 years to less than 10 years
- 10 years to less than 15 years
- 15 years to less than 20 years
- More than 20 years

3. What is your designation?

- Director CEO Senior Executive Other
- If other, Please specify

4. What are the major responsibilities of your position?

5. How often do the senior organisational leadership (e.g. the board of directors) formally meet to discuss/review major issues, strategy and policy?

Part Two: Specific Questionnaire Items

6 (i) Do you have a formal mission statement? Yes No

(ii) If yes, what is it?

7. Who are the major stakeholders* of your organisation? Please mention them in rank order with (i) being the highest rank (for equal priority use the word "and" to join more than one stakeholder).

8. What does each stakeholder identified above expect from your organisation?

9. How do you take into account stakeholder expectations in your strategic planning process?

*See glossary

10. What progress do you want to see in your organisation in 5+ years' time?

11. What high level goals has your organisation established to ensure that it achieves its mission?

12. What strategic objectives has your organisation established in order to measure/monitor progress towards its high level goals? Attempt to list as many strategic objectives as necessary (it is not uncommon to have as many as ten or even more strategic objectives for an organisation, for the purpose of performance monitoring).

13. What specific strategies do you formulate and implement to ensure that your organisation remains economically sustainable? In particular, how do you ensure that donors/agencies remain interested in providing funding to your organisation?

14. List the processes^{*} that your organisation has put in place to achieve each of the strategic objectives that you identified in 12 and 13 above (List by each strategic objective)

Strategic objective	Processes
	Processes_Row_8

*See glossary

15. How does your organisation check/ensure that the strategies that it implements are effective?

16. How do you make your paid employees and volunteers aware that their efforts contribute to achieving the strategic objectives of the organisation?

17. What are the strengths of your organisation that enables it to achieve its strategic objectives?

18. Are you currently using a performance measurement system to measure organisational performance?

- Yes
- No, but planning to use
- No, not planning to use

If the answer to the above question is "yes" is it an in-house performance measurement system or a published performance measurement system (e.g. the balanced score card invented by Robert Kaplan and David Norton, the performance prism invented by Andy Neely)?

- In house
- Published please state the name of the system
- Other please state the name of the system

19. What are the most important performance measures for your organisation? (List as many as necessary)

20 (i) Are you able to identify limitations in some of the performance measures (or the performance measurement system as a whole) that you currently use?

- Yes
- No

(ii) If yes, what are they?

Measure	Noticed Limitation

Measure	Noticed Limitation

21. How does your organisation fulfil its social responsibility?

22. How does your organisation verify/validate that it delivers the expected social impact to the target community? (e.g. actual assessment of change in subjective quality of life over time)

I appreciate your time and effort in providing me the responses.

THANK YOU FOR PARTICIPATING IN THIS STUDY.

If you are willing to participate in a face to face interview or a telephone interview (just in case the researcher needs clarification on some of the responses) which may take up to 1hour at most, please tick the box below to indicate your preferred mode of communication.

Face to face

Telephone

Additional Comments:

APPENDIX B: INVITATION (EMAIL) TO PARTICIPANTS: QUANTITATIVE QUESTIONNAIRE



MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

Dear Sir/Madam,

A Questionnaire to Better Understand the Performance Measurement Dimensions for Australasian Nonprofit Healthcare Organisations

I am Mrs Ishani Soysa, a doctoral student (PhD) from Massey University, Palmerston North, New Zealand. I'm writing to you because you are identified as a valuable person who can provide me with important data for my PhD research project, by way of responding to my questionnaire. The link to my questionnaire is shown at the end of this email.

My aim is to develop and test a performance measurement framework that can be easily adopted by nonprofit healthcare organisations, given their missions, strategic goals and objectives. The questionnaire is a major component of my PhD as it collects data to test my performance measurement framework empirically. I am collecting data from directors and senior executives in Australia and New Zealand, who are responsible for the governance of nonprofit healthcare organisations.

All the data collected from this questionnaire will be used for my PhD research only and the anonymity of the respondents and their organisations will be preserved at all times. This is why I am using Google Forms, which will ensure that I only receive the responses, but not the details of the respondent.

I am more than happy to send an executive summary on the findings of this survey to you. The executive summary will be ready by the end of the year (2015). Should you require the executive summary, do not hesitate to contact me via email at your convenience.

This research is conducted in accordance with Massey University Human Ethics Committee guidelines on low risk research. If you have any queries on completing the questionnaire do not hesitate to contact me.

I would be grateful if you could support my research by way of responding to my quantitative questionnaire, within the next 2 weeks.

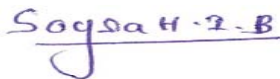
Please visit the following link to response to the questionnaire.

https://docs.google.com/forms/d/1iCod4I1MLcnmKBCGfbq-V3ZVEhYRkUKmmGqE16BZ8YA/viewform?usp=send_form

Thanking you in advance for your support.

Yours faithfully,

Ishani Soysa



BSc (Hons), MSc (Hons)

Doctoral Researcher

School Of Engineering & Advanced Technology

Massey University

Email: I.Soyasa@Massey.ac.nz

Tel: (06)3569099, Ext 85696



MASSEY UNIVERSITY - THE ENGINE OF THE NEW NEW ZEALAND.

New Zealand is emerging with a new confidence on a global stage. As a country we have the desire, innovation and creativity to take on the best the world has to offer. Through our students, staff and alumni, Massey is the engine that is making a practical contribution to turning the wheels in this exciting new chapter in New Zealand. <http://www.engine.ac.nz>

APPENDIX C: THE ONLINE SURVEY QUESTIONNAIRE

A Questionnaire to Better Understand the Performance Measurement Dimensions for Australasian Nonprofit Healthcare Organisations

All the data collected from this questionnaire will be used for my PhD research only and the anonymity of the respondents and their organisations will be preserved at all times. This is why I am using Google Forms, which will ensure that I only receive the responses, but not the details of the respondent.

Please look through the questions carefully, and attempt to answer all the items. It will take about 25-30 minutes of your valuable time to answer all the questionnaire items. The questionnaire consists of two parts. Part One of the questionnaire covers the general information about the respondent and their organisation while Part Two covers 42 statements related to overall organisational behaviour. I am seeking your level of agreement – which can range anywhere from “strongly disagree (=1)” to “strongly agree (=7)”, depending on your perception – to each of these statements.

Part One: Basic information and background about you and your organisation

Please click only one box for each question below

1. Which category mentioned below does your organisation belong to?

- Hospitals and rehabilitation
- Mental health and crisis intervention
- Nursing homes
- Support and ancillary services
- Other:

2. How long has your organisation (including any predecessor/s if applicable) been in operation?

- Less than 5 years
- 5 years to less than 10 years
- 10 years to less than 15 years
- 15 years to less than 20 years
- More than 20 years

3. What is your designation?

- Director
- CEO
- Senior Executive
- Other:

4. In Which country does your organisation based?

- Australia
- New Zealand

5. How familiar are you with performance measurement systems for strategic planning?

- Very Much
- Somewhat
- Very Little

6. Does your organisation use a performance measurement system for strategic planning and/ or for performance monitoring?

- Yes
- No
- Unsure

If you answered 'no' or 'unsure', go to question 8. If 'yes', go to question 7

7. If the answer to question 6 above is yes, for how long have you been using a performance measurement system?

- More than 15 years
- 10 to 15 years
- 5 to 9 years
- Less than 5 years
- Not applicable

8. How many paid employees work in your organisation?

9. Approximately what percentage of funding do you receive from the government and other sources?

(Note: Other category may include a variety of sources and total should equal to 100 Eg. Government - 60% and Other - 40%)

Part Two: Questions relating to your Performance Measurement System

Please indicate your level of agreement by clicking the most appropriate box to each statement in this section (1= Strongly disagree, 2 = Disagree, 3 = Somewhat disagree, 4 = Neither disagree nor agree, 5 = Somewhat agree, 6 = Agree, and 7 = Strongly agree)

1 We support and provide community and social services that effectively meet the needs of those we serve.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

2. We invest in our infrastructure and IT.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

3. We maintain good relationships with the key stakeholders (e.g. funders, contractors, sponsors, and donors) from whom we get our current and future revenue.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

4. We organise projects to raise funds.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

5. We strive to create independence, dignity and unlimited opportunities for our community (or targeted group) to improve their living.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

6. We attempt to measure the quality of life of our clients.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

7. We use recognition schemes to reward our employees for their performance.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

8. We review our services with our clients to collect their feedback.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

9. Our organisation is the recognised leader in all matters related to our targeted group.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

10. We evaluate and improve our processes continuously to achieve better outcomes for our clients.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

11. We collaborate with other organisations to fulfil our social responsibility.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

12. We provide regular information and feedback to the key stakeholders from whom we get our current revenue.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

13. We develop our staff to provide services to our clients efficiently and effectively.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

14. We are continuously innovating new services so as to create new contracts and new streams of revenue.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

15. We constantly endeavour to reduce direct costs and overheads.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

16. Our strategies are formulated based on the requirements of the stakeholders, the environment and the community needs.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

17. Our senior leadership team and the board of directors are significantly involved in planning the strategies for the organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

18. Our strategies are implemented by identifying the resources and the organisational structure needed to implement our strategies.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

19. Business oriented decision making, meaning outcome oriented decision making is important for the sustainability of our organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

20. "Brand awareness" and growing knowledge and interest of the public about the services provided by our organisation is important for the sustainability of our organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

21. We provide fair remuneration to our employees.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

22. We develop and implement new models of services.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

23. We develop, review and update our strategies periodically based on stakeholder feedback, and the state of the environment.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

24. We manage our partnering and collaboration processes thoroughly to ensure the achievement of social responsibility.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

25. We design our service delivery processes to satisfy our stakeholders' requirements.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

26. We use internal and/or external audits to check the quality of services that we are offering.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

27. We design our work/services around the needs of the community or the targeted group.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

28. We invest in our employees regularly (e.g. Benefits, trainings, etc.).

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

29. The ability of the senior leadership team including the board of directors gives strength to our organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

30. We provide training to our staff to maintain a high performance work culture.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

31. We developed organisational infrastructure and technology to deliver quality services to our clients.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

32. We measure the satisfaction of our clients on regular basis.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

33. We measure the satisfaction of our employees on regular basis.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

34. We have a fully integrated IT platform and/or a communication infrastructure.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

35. Our strategies are implemented by identifying the key processes required to implement each strategy.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

36. We develop, review and update our strategies periodically based on performance measurement.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

37. Our key support processes such as HR, Finance, and IT support are designed to support the key service processes.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

38. We hold staff workshops on strategic and business plans to get them involved in the business affairs of our organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

39. Individual and departmental operating plans are traced back to each employee to enable the organisation and the employee see the link between his/her work and the strategic goals of the organisation.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

40. We evaluate and improve our processes continuously to achieve better outcomes for our funders, contractors, sponsors, and donors.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

41. We invest in the development of our volunteers.

1 2 3 4 5 6 7

Strongly Disagree Strongly Agree

I appreciate your time and effort in providing me the responses.

THANK YOU FOR PARTICIPATING IN THIS STUDY.

APPENDIX D: MASSEY UNIVERSITY HUMAN ETHICS APPROVAL – LOW RISK



MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

3 September 2013

Ishani Soysa
3/288 College Street
West End
PALMERSTON NORTH 4410

Dear Ishani

Re: Performance Measurement in Australasian Non-Profit Organisations

Thank you for your Low Risk Notification which was received on 12 August 2013.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

A handwritten signature in black ink that reads "J. O'Neill".

John G O'Neill (Professor)
Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)

cc Dr Nihal Jayamaha
School of Engineering and Advanced Technology
PN321

Assoc Prof Nigel Grigg
School of Engineering and Advanced Technology
PN321

Prof Don Cleland, HoS
School of Engineering and Advanced Technology
PN456

Massey University Human Ethics Committee
Accredited by the Health Research Council

Research Ethics Office

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 350 5573 +64 6 350 5575 F +64 6 350 5622
E humanethics@massey.ac.nz animalethics@massey.ac.nz gtc@massey.ac.nz www.massey.ac.nz

APPENDIX E: QUALITY ASSESSMENT OF THE ANCILLARY MODEL

This appendix shows the results on ‘scale reliability’ (internal consistency reliability) and ‘construct validity’ in respect of the constructs belonging to the ancillary model. The results indicate that the constructs are reliable (in terms of internal consistency) and valid (in terms of convergent and discriminant validities).

Table E1: The Reliability Statistics of the Constructs

Construct	AVE	Composite reliability coefficient (ρ)	Cronbach’s α
Capability	0.75	0.86	0.67
Financial Health	0.82	0.90	0.78
Overall Performance Index	0.63	0.96	0.95
Mission	0.71	0.83	0.61
Stakeholder Satisfaction	0.72	0.89	0.81
Processes	0.92	0.96	0.91
Strategy	0.87	0.93	0.86

Observations: All AVE values > 0.50; all ρ values > 0.70 and all Cronbach’s α > 0.60;

Conclusion: The constructs (more technically, the scales) are reliable

Table E2: Latent Variable Correlations for Discriminant Validity

Construct	Capability	Financial Health	Mission	Stakeholder Satisfaction	Processes	Strategy
Capability	0.87	0	0	0	0	0
Financial Health	0.73	0.90	0	0	0	0
Mission	0.58	0.69	0.84	0	0	0
Stakeholder Satisfaction	0.76	0.80	0.76	0.85	0	0
Processes	0.77	0.82	0.65	0.84	0.96	0
Strategy	0.73	0.75	0.70	0.75	0.79	0.93
Note: The diagonal elements represent the square root of the AVE values of the constructs.						

Observations: High values in the diagonal elements indicate convergent validity (these figures show that the measures are highly correlated with their assigned construct) and the pattern of correlations relative to the figure in the diagonal elements show that the measures show discriminant validity, based on the “Fornell and Larcker criterion” (Fornell & Larcker, 1981).

Conclusion: The measures seem to be producing valid constructs, based on the “Fornell and Larcker Criterion” (Construct is being shown).

Table E3: Loadings and Cross-Loadings for the Ancillary Model

Sub-domains (Items)	First Oder Constructs					
	Capability	Financial Health	Mission	Stakeholder Satisfaction	Processes	Strategy
People	0.88	0.68	0.55	0.67	0.72	0.67
Organisational Infrastructure	0.85	0.59	0.45	0.64	0.61	0.60
Cost Reduction	0.62	0.89	0.56	0.66	0.74	0.64
Budget Increase	0.71	0.91	0.69	0.77	0.74	0.71
Social Profit	0.34	0.44	0.78	0.52	0.44	0.47
Organisational & People Excellence	0.60	0.70	0.90	0.74	0.64	0.69
People Satisfaction	0.69	0.67	0.61	0.84	0.74	0.60
Donor Satisfaction	0.67	0.71	0.68	0.85	0.71	0.75
Client Satisfaction	0.57	0.65	0.64	0.85	0.68	0.56
Process Design	0.74	0.78	0.60	0.78	0.96	0.74
Continuous Improvement	0.73	0.79	0.65	0.82	0.96	0.78
Strategy Implementation	0.68	0.74	0.67	0.75	0.79	0.94
Strategy Planning	0.69	0.65	0.64	0.66	0.68	0.93

Observations: The lodgings are high (> 0.70), showing convergent validity; loadings > cross-loadings, suggesting discriminant validity.

Conclusion: The measures seem to be poducing valid constructs based on observed correlations (Construct Validity is being shown).

References:

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1), 39-50.

**APPENDIX F: THE AUTHOR'S PUBLICATIONS RELATED TO HER
DOCTORAL RESEARCH: THE DRC 16 FORMS**

DRC 16



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

**STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS**

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2016b). Operationalising performance dimensions for the Australasian nonprofit healthcare sector. *TQM Journal*, 28(6), forthcoming.

In which Chapter is the Published Work: Chapter Six

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 85% and / or
- Describe the contribution that the candidate has made to the Published Work:
Lit review, methodology, entire fieldwork, results and discussion were solely candidate's contributions; so is the writing; the first draft of the paper and revisions in the second round of the review of the paper were done by the candidate. Supervisors (the coauthors) were playing only a supporting role (e.g. methods verification, few tidying up work), as advisors.

Ishani Soysa
Digitally signed by Ishani Soysa
DN: cn=Ishani Soysa, o=Massey University,
ou=GRS, email=IshaniSoysa@massey.ac.nz,
c=NZ
Date: 2016.07.21 16:48:02 +1200

Candidate's Signature

21 July 2016

Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
DN: cn=Nihal Jayamaha, o=Massey University,
ou=School of Engineering and Advanced Technology,
email=N.Jayamaha@massey.ac.nz, c=NZ
Date: 2016.07.21 16:17:02 +1200

Principal Supervisor's signature

21 July 2016

Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2016). Validating the Balanced Scorecard framework for nonprofit organisations: An empirical study involving Australasian healthcare. *Total Quality Management and Business Excellence*, forthcoming (see note below).

In which Chapter is the Published Work: Chapter Eight and Nine

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 85%
and / or
- Describe the contribution that the candidate has made to the Published Work:

The above paper is in the final round of revision (minor revisions) and the authors are very confident that the paper will get accepted very soon and will appear online within 2016. The candidate conducted the lit review, designed the entire methodology (all fieldwork included), results and discussion and the conclusions; so is the writing; the first draft of the paper and revisions in the second round of review of the paper were done by the candidate. Supervisors (the coauthors) were playing only a supporting role (e.g. methods verification, few tidying up work), as advisors.

Ishani Soysa
Digitally signed by Ishani Soysa
DN: cn=Ishani Soysa, o=Massey University,
ou=GRS, email=I.Soyas@massey.ac.nz,
c=NZ
Date: 2016.07.21 17:13:46 +1200

Candidate's Signature

21 July 2016
Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
DN: cn=Nihal Jayamaha, o=Massey University,
ou=GRS, email=N.Jayama@massey.ac.nz,
c=NZ
Date: 2016.07.21 17:06:10 +1200

Principal Supervisor's signature

21 July 2016
Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2016a, 26-29 June). Developing a performance scoring system for nonprofit organisations. Paper presented at the 10th conference of the Performance Measurement Association, Edinburgh, Scotland.

In which Chapter is the Published Work: Chapter Eight and Nine

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 90% and / or
- Describe the contribution that the candidate has made to the Published Work:
All aspects of the paper from the introduction, lit review, methodology, fieldwork, results and discussion through to conclusion were solely candidate's contributions. Supervisors (the coauthors) were playing only a supporting role with minor tidying up of the paper in the draft version and the revised version.

Ishani Soysa
Digitally signed by Ishani Soysa
 DN: cn=Ishani Soysa, o=Massey University,
 ou=GRS, email=I.Soyasa@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:49:28 +1200

Candidate's Signature

21 July 2016

 Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
 DN: cn=Nihal Jayamaha, o=Massey University,
 ou=School of Engineering and Advanced Technology,
 email=N.Jayamaha@massey.ac.nz, c=NZ
 Date: 2016.07.21 16:53:07 +1200

Principal Supervisor's signature

21 July 2016

 Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

**STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS**

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2015, 31 May - 2 June). Operationalising performance measurement in the Australasian nonprofit healthcare sector - An empirical study. Paper presented at the 13th ANZAM Operations, Supply Chain and Services Management Symposium, RMIT University, Melbourne.

In which Chapter is the Published Work: Chapter Six and Nine

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 90%
and / or

- Describe the contribution that the candidate has made to the Published Work:

All aspects of the paper from the introduction, lit review, methodology, fieldwork, results and discussion through to conclusion were solely candidate's contributions. Supervisors (the coauthors) were playing only a supporting role with minor tidying up of the paper in the draft version and the revised version.

Ishani Soysa
Digitally signed by Ishani Soysa
 DN: cn=Ishani Soysa, o=Massey University,
 ou=GRS, email=Ishani.Soyas@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:01:00 +1200

Candidate's Signature

21 July 2016

 Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
 DN: cn=Nihal Jayamaha, o=Massey University,
 ou=School of Engineering and Advanced Technology,
 email=N.Jayamaha@massey.ac.nz, c=NZ
 Date: 2016.07.21 16:13:14 +1200

Principal Supervisor's signature

21 July 2016

 Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

**STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS**

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2014, 3 - 4 July). Developing a conceptual model for measuring performance of Australasian nonprofit sector. Paper presented at the 12th ANZAM Operations, Supply Chain and Services Management Symposium, University of Auckland Business School.

In which Chapter is the Published Work: Chapter Four

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 90% and / or

- Describe the contribution that the candidate has made to the Published Work:

All aspects of the paper from the introduction, lit review, methodology, discussion through to conclusion were solely candidate's contributions. Supervisors (the coauthors) were playing only a supporting role with minor tidying up of the paper in the draft version and the revised version.

Ishani Soysa
Digitally signed by Ishani Soysa
 DN: cn=Ishani Soysa, o=Massey University,
 ou=GRS, email=I.Soyas@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:50:50 +1200

Candidate's Signature

21 July 2016

 Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
 DN: cn=Nihal Jayamaha, o=Massey University,
 ou=School of Engineering and Informal Technology,
 email=N.Jayamaha@massey.ac.nz, c=NZ
 Date: 2016.07.21 16:55:50 +1200

Principal Supervisor's signature

21 July 2016

 Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

**STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS**

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soysa

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soysa, I. B., Jayamaha, N. P., & Grigg, N. P. (2013b, 20-21 June). Performance measurement in the nonprofit sector. Paper presented at the 11th ANZAM Operations, Supply Chain and Services Management Symposium, Research Industry Partnership Brisbane.

In which Chapter is the Published Work: Chapter Three

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 85% and / or
- Describe the contribution that the candidate has made to the Published Work:

All aspects of the paper from the introduction, lit review, to research questions to conclusion were solely candidate's contributions. Supervisors (the coauthors) were playing only a supporting role with minor tidying up of the paper in the draft version and the revised version. Note that this paper was written as a review article to identify knowledge gaps.

Ishani Soysa
Digitally signed by Ishani Soysa
 DN: cn=Ishani Soysa, o=Massey University,
 ou=GRS, email=I.Soyas@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:54:56 +1200

Candidate's Signature

21 July 2016

Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
 DN: cn=Nihal Jayamaha, o=Massey University,
 ou=School of Engineering and Informal Technology,
 email=N.Jayamaha@massey.ac.nz, c=NZ
 Date: 2016.07.21 16:33:06 +1200

Principal Supervisor's signature

21 July 2016

Date



MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: H. Ishani Buddika Soya

Name/Title of Principal Supervisor: Dr Nihal Jayamaha

Name of Published Research Output and full reference:

Soya, I. B., Jayamaha, N. P., & Grigg, N. P. (2013a, 30 October- 1 November). Examining the performance measurement in the Australasian nonprofit sector. Paper presented at the Performance Management Association of Australasia (PMAA) Conference 2013, Queenstown, New Zealand.

In which Chapter is the Published Work: Chapter Two and Three

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 95% and / or
- Describe the contribution that the candidate has made to the Published Work:

The full paper. This paper was an updated version of the previous review article with more refined research questions.

Ishani Soya
Digitally signed by Ishani Soya
 DN: cn=Ishani Soya, o=Massey University,
 ou=GRS, email=I.Soya@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:58:50 +1200

Candidate's Signature

21 July 2016

 Date

Nihal Jayamaha
Digitally signed by Nihal Jayamaha
 DN: cn=Nihal Jayamaha, o=Massey University,
 ou=School of Engineering and Applied Technology,
 email=N.Jayamaha@massey.ac.nz,
 c=NZ
 Date: 2016.07.21 16:38:00 +1200

Principal Supervisor's signature

21 July 2016

 Date



Emerald Insight



The TQM Journal

Operationalising performance measurement dimensions for the Australasian nonprofit healthcare sector

Ishani Buddika Soysa Nihal Palitha Jayamaha Nigel Peter Grigg

Article information:

To cite this document:

Ishani Buddika Soysa Nihal Palitha Jayamaha Nigel Peter Grigg , (2016), "Operationalising performance measurement dimensions for the Australasian nonprofit healthcare sector", The TQM Journal, Vol. 28 Iss 6 pp. 954 - 973

Permanent link to this document:

<http://dx.doi.org/10.1108/TQM-08-2015-0109>

Downloaded on: 30 October 2016, At: 18:38 (PT)

References: this document contains references to 70 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 75 times since 2016*

Users who downloaded this article also downloaded:

(2016), "The mediating role of TQM and organizational excellence, and the moderating effect of entrepreneurial organizational culture on the relationship between ERP and organizational performance", The TQM Journal, Vol. 28 Iss 6 pp. 991-1011 <http://dx.doi.org/10.1108/TQM-04-2014-0040>

(2016), "Using HCAHPS data to improve hospital care quality", The TQM Journal, Vol. 28 Iss 6 pp. 974-990 <http://dx.doi.org/10.1108/TQM-09-2014-0072>

Access to this document was granted through an Emerald subscription provided by emerald-sm:191620 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Operationalising performance measurement dimensions for the Australasian nonprofit healthcare sector

Ishani Buddika Soysa, Nihal Palitha Jayamaha and
Nigel Peter Grigg

*School of Engineering and Advanced Technology, Massey University,
Palmerston North, New Zealand*

Abstract

Purpose – The purpose of this paper is to develop a performance measurement (PM) framework for Australasian nonprofit organisations (NPOs) involved in healthcare, and operational descriptions for each PM dimension within this framework.

Design/methodology/approach – Literature relating to the balanced scorecard and other PM frameworks was examined to develop an initial conceptual model, and this model was substantially improved by collecting qualitative data from nine Australian and New Zealand healthcare NPOs using a case study approach.

Findings – The study identifies nine causally related PM dimensions: mission, strategy, organisational capabilities, infrastructure and people development (people and information), financial health, processes, and stakeholder satisfaction (clients, people, and donors). The study also recognised that “Mission” and “Strategy” should be PM dimensions and that healthcare NPOs should focus on satisfying its people, not only donors and clients. Additionally, 41 operational descriptions are developed for each of these dimensions and can enable detailed PM items to be derived by organisations.

Originality/value – The study is the first study that has been undertaken to develop a PM framework for the Australasian NPOs to a level that it can be readily used by the practitioners (following customisation to their own specific context). The developed model also serves as a basis for future quantitative academic research aimed at testing and empirical validation of the conceptual model.

Keywords Performance measurement, Case study, Balanced scorecard, Nonprofit sector

Paper type Research paper

1. Introduction

All organisations exist to achieve certain high-level objectives. In the case of nonprofit organisation (NPO), most of the objectives relate to serving the target communities to make some positive impact on them. No matter to which sector an organisation belongs (for-profit, public, nonprofit), it needs to measure how it is performing in achieving its strategic objectives (Douglas *et al.*, 2010; Fryer *et al.*, 2007). A performance measurement (PM) system can be defined as a “set of metrics used to quantify both the efficiency and effectiveness of actions” implemented to achieve the strategic goals of an organisation (Neely *et al.*, 2005, p. 1229). The theoretical framework of a PM system facilitates the generation of performance metrics along predefined measurement dimensions. For each PM dimension, managers need to determine the appropriate set of measures to gauge the organisational performance (Kennerley and Neely, 2002) and the measures themselves should possess the following features: they should be derived from the strategy (Anthony and Govindarajan, 2007; Neely *et al.*, 2005), they, as a whole, should be capable of providing a balanced view of performance (Kaplan and





Validating the Balanced Scorecard Framework for Nonprofit Organisations: An Empirical Study Involving Australasian Healthcare

Journal:	<i>Total Quality Management & Business Excellence</i>
Manuscript ID	CTQM-2015-0337.R1
Manuscript Type:	Original Article
Keywords:	Performance measurement , Balanced Scorecard , Structural equation modelling , Nonprofit healthcare organisations, Australasia
Abstract:	In this paper we theorise and empirically validate the nonprofit version of the balanced scorecard (BSC) using Australasian nonprofit healthcare organisations (NPHOs) as the sampling frame. We used 232 valid responses (collected via a questionnaire that captures the conceptual domain of the 9 constructs of our model) to test our hypotheses using partial least squares based structural equation modelling (PLSBSEM). Our study brings clarity to the nonprofit BSC as we explain how organisational Mission drives the Strategy (which in turn drives the system) to achieve the stakeholder (Client, People, and the Donors) outcomes. Our hypotheses (which are ground in the literature) were supported by the data and we discussed the implications of the findings from a theoretical and practical standpoint to guide future research.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

Validating the Balanced Scorecard Framework for Nonprofit Organisations: An Empirical Study Involving Australasian Healthcare

Abstract

In this paper we theorise and empirically validate the nonprofit version of the balanced scorecard (BSC) using Australasian nonprofit healthcare organisations (NPHOs) as the sampling frame. We used 232 valid responses (collected via a questionnaire that captures the conceptual domain of the 9 constructs of our model) to test our hypotheses using partial least squares based structural equation modelling (PLSBSEM). Our study brings clarity to the nonprofit BSC as we explain how organisational Mission drives the Strategy (which in turn drives the system) to achieve the stakeholder (Client, People, and the Donors) outcomes. Our hypotheses (which are ground in the literature) were supported by the data and we discussed the implications of the findings from a theoretical and practical standpoint to guide future research.

Key words: Performance measurement (PM), Balanced Scorecard (BSC), Nonprofit healthcare organisations (NPHOs), Australasia, Structural equation modelling (SEM)

1. Introduction

It is hard to know in which direction an organisation is heading, unless its goals become measurable. Therefore, no matter to which sector an organisation belongs—private, public, or nonprofit—it is useful to have a robust measurement system to gauge how it is performing in achieving its goals. Managers use a performance measurement (PM) system such as the balanced scorecard (BSC) to develop performance measures to suit their organisations to examine the efficiency and effectiveness of the organisational strategies as well as to convey