Calls for Service and Police Effectiveness: The Role of Performance Measurement Systems

Gavin Mark Knight
BE(Hons) DipORS MSc(Dist)

School of Criminology and Criminal Justice
Arts, Education and Law
Griffith University

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy
August 2014
Abstract

This thesis explores the adoption of public sector management reforms by police. The literature suggests that police have been more resistant to public sector management reform than other public sector organisations. One feature of public sector management reform—performance measurement—is evident in police agencies, but prior research does not provide a clear picture of whether performance measurement helps or hinders police effectiveness. The research presented in this thesis helps to clarify this picture by examining the influence of performance measurement on the effectiveness of police in responding to calls for service from citizens.

Responding to calls for service involves managing numerous multi-faceted challenges. Performance measurement can help manage complexity by enabling members of an organisation to make sense of what is happening in complex organisational contexts. Therefore, the context of police response to calls for service from citizens is one where performance measurement can assist police to be effective. In this thesis I examine whether performance measurement does assist police to be effective in responding to calls for service.

The research presented in this thesis addresses four aims:

Aim1. examine what police understand effectiveness in responding to calls for service to be;

Aim 2. determine what influence performance measurement has on police effectiveness;
Aim 3. explore why performance measurement has the influence it does on police effectiveness; and

Aim 4. explore the influence of performance measurement on the perspectives and behaviours of police personnel.

A mixed methods research design is used in this thesis to address these aims and examine the relationship between performance measurement and police effectiveness in responding to calls for service. The design includes two interconnected research studies: The first study is exploratory because prior research on performance measurement contains gaps in the perspectives and behaviours of police personnel in relation to using performance measurement. Through qualitative interviews with 121 police personnel from New Zealand Police and Queensland Police Service, I investigate the perspectives and behaviours of personnel in these police agencies who deliver and manage a service of responding to calls for service from citizens. Accounts from these personnel reveal that police have a good understanding of various dimensions of police effectiveness in responding to calls for service. However, tension between technical-rational and institutional goals creates a loosely coupled environment in which barriers exist to police making effective use of performance measurement.

For Study 2 I draw on the findings of Study 1 and the literature to construct four hypotheses about the relationship between performance measurement and police effectiveness in responding to calls for service from citizens. These hypotheses are:

Hypothesis 1. The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.
Hypothesis 2. Police do not respond to performance measures in a way that influences their organisational legitimacy.

Hypothesis 3. Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.

Hypothesis 4. Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

I test these hypotheses using administrative data from New Zealand Police and Queensland Police Service and transcripts of New Zealand parliamentary debates. I use transfer function analysis to examine the hypothesised relationships between variables. Corroborating findings from Study 1, I identify limitations in how police use performance information. I conclude that performance measurement can have a positive influence on police effectiveness. However, this requires a different approach than most police managers currently use in practice. I suggest that by embedding analytical capability in police organisations and by increasing external scrutiny, police might stimulate more effective use of performance information.
Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Gavin Knight
Acknowledgements

I must confess that I set out to gain a qualification, but ended up gaining an education. For this I must thank my supervisors, Professor Anna Stewart, Professor Lorraine Mazerolle, and Dr Jacqueline Drew. I appreciate the challenge they faced in shaping someone with a background in engineering and mathematics into a social scientist. They taught me to think in new ways and appreciate new disciplines in how to research and write. I feel very fortunate that such dedicated and skilful people were willing to invest of themselves in me. I will forever be in their debt.

I am also grateful to Professor David McDowall of State University of New York, and Mr Alistair Gray of Statistics Research Associates Limited, who provided technical advice and training on selection and application of the transfer function analysis method I employed in Chapters 6 and 7. Also, thanks to Dr Dalice Sim of Victoria University of Wellington, who checked the analysis I undertook in Chapters 6 and 7.

Importantly, I would like to thank Queensland Police Service and New Zealand Police. These organisations assisted me through the provision of data and access to their personnel. The findings of this thesis are based, in large part, on data supplied by these organisations and interviews with personnel from these organisations. However, I acknowledge that this does not mean these findings necessarily represent the views of Queensland Police Service or New Zealand Police. Any errors or omissions are the responsibility of the author.

Next I would like to thank the many personnel of Queensland Police Service and New Zealand Police who were so open and willing with their time and support of my research.
I would also like to thank the New Zealand Police for supporting my professional development.

Last I would like to acknowledge my family. In particular, my daughters Rebecca and Tabitha supported me with some of the tedious mechanical aspects of processing data, and my wife Glenys put up with my working many late nights and traveling away to collect data on numerous occasions; all this through a period when our family had to cope with a number of significant stresses. Thank you.
# Contents

Abstract .............................................................................................................................................. 1

Statement of Originality ................................................................................................................ 4

Acknowledgements ......................................................................................................................... 5

Contents ............................................................................................................................................ 7

List of tables .................................................................................................................................... 13

List of figures .................................................................................................................................. 14

Glossary ............................................................................................................................................ 15

Chapter 1: Research Agenda ......................................................................................................... 18
  1.1 Introduction ............................................................................................................................... 18
  1.2 The Scope of This Thesis .......................................................................................................... 20
  1.3 Moving Beyond Prior Research .............................................................................................. 22
  1.4 Thesis Outline .......................................................................................................................... 25

Chapter 2: Performance Measurement in Police ............................................................................. 26
  2.1 Introduction ............................................................................................................................... 26
  2.2 The History and Nature of New Public Management .............................................................. 27
      2.2.1 The origin of New Public Management .......................................................................... 27
      2.2.2 Public choice theory and principal-agent theory ......................................................... 29
      2.2.3 Does NPM improve effectiveness? .................................................................................. 30
  2.3 Factors That Influence an Organisation’s Use of Performance Information..................... 33
  2.4 The Adoption of NPM by Police ............................................................................................... 35
2.4.1 Technical-rational theory and institutional theory ........................................... 37

2.4.2 Compstat ........................................................................................................... 43

2.4.3 The evolution of NPM in police around the world ........................................... 45

2.5 Why Measure Performance at all? ........................................................................ 49

2.6 Conclusion ............................................................................................................ 51

Chapter 3: Police Response to Calls for Service ......................................................... 53

3.1 Introduction ............................................................................................................ 53

3.2 The Challenges Police Face in Responding Effectively to Calls for Service ...... 55

  3.2.1 Citizens reporting to police ............................................................................. 56

  3.2.2 Handling calls from the public ......................................................................... 59

  3.2.3 Dispatch ......................................................................................................... 61

  3.2.4 Attendance at the scene of the call out ......................................................... 65

  3.2.5 The impact of technological innovation ............................................................ 70

3.3 Dimensions of Police Effectiveness ....................................................................... 73

3.4 Conclusion ............................................................................................................ 77

Chapter 4: Study 1 Method .......................................................................................... 80

4.1 Research Questions ............................................................................................... 80

4.2 Interview Sites ...................................................................................................... 83

4.3 Participants ........................................................................................................... 86

  4.3.1 Ranks and roles of participants ..................................................................... 86

  4.3.2 Procedure used to select participants .............................................................. 91
4.4 Semi-structured Interviews .................................................................................. 94

4.5 Analytic Approach ............................................................................................... 100

Chapter 5. Study 1 Findings ...................................................................................... 102

5.1 Introduction ........................................................................................................... 102

5.2 RQ1 How do Police Managers use Performance Information, and do They use Performance Information Effectively? ......................................................... 103

5.2.1 RQ1.1 What types of performance information do police managers use? 104

5.2.2 RQ1.2 Why do police managers make only limited use of quantitative performance information? ................................................................. 106

5.2.3 RQ1.3 How do police managers use quantitative performance information to influence decisions? ................................................................. 108

5.3 RQ 2 Does Performance Measurement Help or Hinder Police Effectiveness in Either Technical-rational or Institutional Terms? ............................................. 112

5.3.1 RQ2.1 What do police personnel believe is the reason for police measuring their performance in responding to calls for service? ............................... 115

5.3.2 RQ2.2 Do police personnel perceive performance measurement in police to be helpful? ........................................................................................................ 119

5.3.3 RQ2.3 What performance measurement mechanisms do police use? ...... 123

5.3.4 RQ2.4 What performance measures do police monitor? .............................. 135

5.4 RQ3 What Influence Does Performance Measurement in Police Organisations Have on Loose Coupling, and What Influence Does Loose Coupling Have on Police Effectiveness? ...................................................................................... 143
5.4.1 Police Perspectives on What Constitutes Effectiveness in Responding to Calls for Service ............................................................................................................................................... 143

5.4.2 Evidence of loose coupling in New Zealand Police and Queensland Police Service .............................................................................................................................................. 149

5.5 Moving Beyond Prior Research .................................................................................................................................................................................. 158

5.6 Generating Hypotheses for Study 2 ............................................................................................................................................................................. 161

5.6.1 Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service .................... 164

5.6.2 Hypothesis 2: Police do not respond to performance measures in a way that influences their organisational legitimacy ......................................................... 166

5.6.3 Hypothesis 3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service .... 167

5.6.4 Hypothesis 4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take .................................................................................................. 168

5.7 Summary .............................................................................................................................................................................................................. 168

Chapter 6: Study 2 Method ....................................................................................................................................................................................... 170

6.1 Introduction ........................................................................................................................................................................................................ 170

6.2 Selection of Method ......................................................................................................................................................................................... 172

6.3 Step 1: Select Variables ............................................................................................................................................................................... 175

6.4 Step 2: Acquire Data ...................................................................................................................................................................................... 181

6.5 Step 3: Pre-whiten Time Series ................................................................................................................................................................. 186
6.6 Step 4: Cross-correlate Time Series ................................................................. 188

6.7 Step 5: Perform Granger-causality Tests ......................................................... 189

Chapter 7: Study 2 Findings ............................................................................. 191

7.1 Introduction ................................................................................................. 191

7.2 Descriptions of Hypotheses ..................................................................... 192

7.2.1 Hypothesis 1: Performance measurement and police effectiveness .... 192

7.2.2 Hypothesis 2: Performance measurement and institutional effectiveness .. 194

7.2.3 Hypothesis 3: Institutional effectiveness and technical-rational effectiveness .................................................................................................................. 195

7.2.4 Hypothesis 4: Morale and performance .................................................. 196

7.3 Interpreting Results of Hypothesis Tests ...................................................... 198

7.4 Findings ...................................................................................................... 200

7.4.1 Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service ................. 201

7.4.2 Hypothesis 2: Police do not respond to performance measures in a way that influences their organisational legitimacy ......................................................... 205

7.4.3 Hypothesis 3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service .... 206

7.4.4 Hypothesis 4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take .................................................................................................................. 211

7.5 Summary of Findings ................................................................................. 213
Chapter 8: Implications for Theory, Policy, and Practice ................................. 216

8.1 Introduction ........................................................................................................ 216

8.2 Main Findings...................................................................................................... 218

  8.2.1 Breadth of performance measures ............................................................... 219

  8.2.2 The influence of performance measurement on police effectiveness......... 221

  8.2.3 Why police measure performance ............................................................... 222

  8.2.4 Loose coupling ............................................................................................ 223

  8.2.5 Conclusions .................................................................................................. 224

8.3 Implications for Policy and Practice ................................................................. 225

  8.3.1 Argument for preserving the current state ................................................ 226

  8.3.2 Argument for changing the current state .................................................... 228

8.4 Caveats and Opportunities for Future Research .............................................. 229

  8.4.1 Perspectives of citizens .............................................................................. 230

  8.4.2 Limitations of monthly time series data ..................................................... 231

  8.4.3 Limitations of transfer function analysis .................................................... 232

  8.4.4 Generalisability beyond Queensland and New Zealand .............................. 233

8.5 Concluding Comments ....................................................................................... 234

References.............................................................................................................. 236

Appendix A: Informed Consent Package ................................................................. 269

Appendix B: Interview Schedules .......................................................................... 274

Appendix C: Results of Tests Performed in Study 2 .............................................. 293
List of tables

Table 4.1 State and Police Populations ..............................................................84
Table 4.2 Interviewees by Rank .....................................................................90
Table 4.3 Interviewees by Role .....................................................................90
Table 4.4 Summary of Interview Schedules ..................................................96
Table 5.1 Proportion of Interviewees who Referred to each mechanism .........123
Table 5.2 Proportion of Interviewees who Referred to OPRs .........................125
Table 5.3 Proportion of Interviewees who Referred to Individual Performance Appraisals ..........129
Table 5.4 Proportion of Interviewees who Referred to each measure ..........136
Table 5.5 Proportion of Interviewees who Identified Response Time as a Measure ........137
Table 7.1 New Zealand Results for Hypothesis 1 ..........................................202
Table 7.2 Queensland Results for Hypothesis 1 ..........................................203
Table 7.3 New Zealand Results for Hypothesis 2 ......................................205
Table 7.4 Queensland Results for Hypothesis 2 ......................................206
Table 7.5 New Zealand Results for Hypothesis 3 ......................................207
Table 7.6 Queensland Results for Hypothesis 3 ......................................210
Table 7.7 New Zealand Results for Hypothesis 4 ......................................212
Table 7.8 Queensland Results for Hypothesis 4 ......................................212
List of figures

Figure 5.1 The Influence of Performance Measurement on Police Effectiveness ........................................ 162

Figure 7.1 Relationships Between Inputs to and Outcomes of Performance Measurement ...................... 193

Figure 7.2 Relationships Between Staff Headcount and Performance ..................................................... 194

Figure 7.3 Relationships Between Performance Measures and Institutional Effectiveness .................... 195

Figure 7.4 Relationships Between Indicators of Institutional Effectiveness and Technical-rational Effectiveness ................................................................................................................................................. 196

Figure 7.5 Relationships Between Sick Leave and Effectiveness ............................................................. 197

Figure 7.6 Time Series of NV ................................................................................................................... 207

Figure 7.7 Correlogram of ARIMA Pre-whitened NV ............................................................................. 208

Figure 7.8 Correlogram of VAR-AIC Pre-whitened NV .......................................................................... 209

Figure 7.9 Correlogram of VAR-12 Pre-whitened NV ........................................................................... 209
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAD</td>
<td>Computer Aided Dispatch. Queensland Police Service acronym for the dispatch management system.</td>
</tr>
<tr>
<td>CARD</td>
<td>Communications And Resource Deployment. New Zealand Police acronym for the dispatch management system.</td>
</tr>
<tr>
<td>Correlogram</td>
<td>A graphical representation of correlation statistics. The correlogram of a time series is also known as an autocorrelation plot.</td>
</tr>
<tr>
<td>Cross-correlogram</td>
<td>A correlogram presenting correlation statistics for the cross-correlation of two time series</td>
</tr>
<tr>
<td>DPR</td>
<td>Differential Police Response (see Section 3.2.3)</td>
</tr>
<tr>
<td>Hansard</td>
<td>A database containing transcripts of parliamentary debates</td>
</tr>
<tr>
<td>Institutional</td>
<td>A theoretical perspective that views organisational structures as being influenced by cultural features of their environments (see Section 2.4.1)</td>
</tr>
<tr>
<td>Loose-coupling</td>
<td>A theoretical perspective that addresses breakdown in alignment between members of an organisation (see Section 2.4.1)</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Noise</td>
<td>A colloquial term used by some interviewees in Study 1 that refers to the volume of negative commentary in the media or parliament on a particular issue about police</td>
</tr>
<tr>
<td>NPM</td>
<td>New Public Management (see Section 2.2)</td>
</tr>
<tr>
<td>Operations</td>
<td>A mathematical discipline concerned with optimising decisions research</td>
</tr>
<tr>
<td>OPR</td>
<td>Operational Performance Review (Queensland) / Organisational Performance Review (New Zealand). A formal performance measurement and review framework used within police organisations.</td>
</tr>
<tr>
<td>Organisational legitimacy</td>
<td>The approval of society for an organisation (see Section 2.4.1)</td>
</tr>
<tr>
<td>PCC</td>
<td>Police Communications Centre where calls for service are received and units are dispatched to attend jobs</td>
</tr>
<tr>
<td>Performance information</td>
<td>Management information about how well an organisation or part of an organisation is performing. Performance information can take the form of a set of performance indicators, which are quantitative measures of organisational effectiveness or efficiency.</td>
</tr>
</tbody>
</table>
Performance management: A process by which an organisation aligns its resources, activities, and employees to achieve its objectives, typically by acting in response to performance information.

Performance measurement: The systematic monitoring, interpreting, and reporting of performance information.

Technical-rational: A theoretical perspective that characterises an organisation by the products and services it delivers (see Section 2.4.1).

Unit: The police officer or officers dispatched to attend calls for service (also referred to as crew in Queensland).
Chapter 1: Research Agenda

1.1 Introduction

Society expects the police to perform a variety of functions. Governments expect police to maintain order (Bayley, 1994b) and to demonstrate to both offenders and the public that a regime of law exists (Silberman, 1978). Democratic governments also wish to reinforce their legitimacy as a democracy by demonstrating responsiveness to citizens (Herbert, 2006) who expect the police to respond when called (Sparrow, Moore, & Kennedy, 1990). However, to deliver an effective service of responding to calls for service police must manage numerous complicated challenges including illegitimate calls, variable call volumes throughout the day, and conflicting priorities. If police make mistakes this may result in serious consequence to life, property, and the reputation of police (Kuhn & Hoey, 1987; Moskos, 2007; Sparrow et al., 1990). For example, in reviewing the circumstances contributing to the death of Ms Iraena Asher:

Coroner Ryan said contributing to the death was the failure of the police not to send a car to Piha to ascertain whether Ms Asher needed help after she made a distressed 111 call. But instead of sending a patrol car, police called a taxi for her. It never arrived, and it later emerged it had been sent to the wrong address. (Leask, 2012)

Complaints about the police’s performance in this and other incidents in 2004 brought into question the effectiveness of the New Zealand Police emergency response system (B. Carter, 2004a, 2004b; Hansard, 2004; Plowman, 2004). Public confidence in the police was affected. The police commissioner responded by initiating an independent external
review that found numerous systemic deficiencies in police practices and culture, and initiated an extensive programme of reform (Corboy, Gilbert, Purdie, & McKenna, 2005).

Similarly, following the failure of Queensland Police Service to respond to a call for service that might have prevented the death of Ms Cherie Cundy, the effectiveness of the police response was investigated. In reviewing the circumstances contributing to the death of Ms Cundy, a Brisbane Coroner noted that police had insufficient staff available on the night of Ms Cundy’s death to respond to the volume of calls for service police received (Hutton, 2012). As well as the coroner’s investigation, Queensland Police Service investigated its failure to respond to the initial call received, and changed its policy for allocating resources to attend calls for service.

Organisations monitor and shape organisational behaviour (Hoggett, 1996) by measuring and reporting performance indicators to improve rational decision-making (Johnsen, 2005). Both New Zealand Police and Queensland Police Service operate performance measurement frameworks which aim to ensure that police perform effectively. These frameworks seek to detect and remedy systemic problems in advance of crises of organisational legitimacy occurring. However, crises of organisational legitimacy, such as illustrated above, do occur, and trigger changes in police operational policies. It is unclear how well police use performance measurement to detect and remedy systemic problems and perform effectively. In this thesis I examine how police in New Zealand and Queensland use performance measurement, and what influence this has on their effectiveness in responding to calls for service from citizens.
1.2 The Scope of This Thesis

Ensuring delivery of an effective service in the complex and politically charged context in which police operate has proved challenging for police around the world (Kuhn & Hoey, 1987; Moskos, 2007; Sparrow et al., 1990). Throughout the twentieth century police implemented reforms aimed at improving their effectiveness. However, such reforms often produced undesirable consequences or failed to be fully successful (Chan, 2001; Mazerolle, Rogan, Frank, Famega, & Eck, 2002; Reiss, 1992; Sparrow et al., 1990; Weisburd & Eck, 2004). Many police reforms, such as intelligence-led policing, problem-oriented policing, and community-oriented policing, aimed to make police more proactive in reducing crime (Dammer, Fairchild, & Albanese, 2006; Eck & Spelman, 1987; Goldstein, 1990; Greene & Mastrofski, 1988; Ratcliffe, 2003; Telep & Weisburd, 2012). However, citizens expect the police to provide a responsive service as well (Sparrow et al., 1990). Police are regularly called to attend many different types of situations. These include natural disasters, traffic accidents, crimes, public disorder, and numerous different circumstances where someone thinks that a competent authority figure is required (Bayley, 1990, 1994b; Sparrow et al., 1990; Terrill, 2003).

In this thesis I will show that delivering an effective responsive service involves many complexities and challenges for police. Police operate performance measurement frameworks that can help manage this complexity by enabling members of an organisation to make sense of what is happening in complex organisational contexts (Gill, 2011). However, many scholars observe that although police closely monitor measures related to crime they ignore the many other things that police do (Fleming, 2009; Kelling, 1992; Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; M. H. Moore & Braga, 2003; Skogan & Frydl, 2004). These scholars argue that the absence of broader
performance measures limits the effectiveness of police. In this thesis I explore these critical concerns and examine their applicability to police in New Zealand and Queensland.

I also explore how police use performance information, and the influence of performance measurement on police effectiveness in responding to calls for service. To improve the performance of public sector agencies, reforms in the 1980s and 90s implemented performance measurement frameworks (Fleming & Lafferty, 2000; Gill, 2011; Hoggett, 1996). Some scholars (Cope, Leishman, & Starie, 1997) identify police reform since the 1980s as a manifestation of public sector management reform. To understand the influence of performance measurement on police effectiveness I explore the theoretical basis for this public sector reform and evidence of its influence on the public sector in general and police in particular.

Explanations given by technical-rational theory and institutional theory are frequently contrasted in the literature on public sector reform (Collier, 2001; Gruening, 2001; Willis, Mastrofski, & Weisburd, 2007). The technical-rational perspective characterises an organisation by the products and services it delivers (Willis et al., 2007). An organisation’s performance relates to how effectively and efficiently it delivers these products and services. The institutional perspective (Meyer & Rowan, 1977) views organisational structures as being influenced by cultural features of their environments. Rather than being driven to perform according to technical-rational criteria, organisations are judged by how well they respond to beliefs about what they should look like and what they should be doing. Organisations gain legitimacy with their constituents by reflecting these beliefs. Perceptions by an organisation’s employees of misalignment or confusion
between technical-rational goals and institutional goals can lead to what is known as *loose coupling* (Lawton, McKeivitt, & Millar, 2000; Meyer & Rowan, 1977). Loose coupling manifests as a gap between formal structures and work practices. Formal rules are often broken, and decisions are often not implemented (Collier, 2001; P. Smith, 1993).

In this thesis I explore the nature of loose coupling in police, and examine its influence on the formal structures and work practices surrounding performance measurement in police. I explore evidence in the literature on loose coupling. Drawing from the literature, I identify a number of key questions which I seek to answer by gathering evidence on the perspectives of New Zealand and Queensland police personnel who deliver or manage the service of responding to calls for service from citizens. Using insights gained from these perspectives I frame hypotheses about the relationship between performance measurement and police effectiveness in responding to calls for service from citizens. I then use police administrative data to test these hypotheses and extend the understanding provided by prior research.

### 1.3 Moving Beyond Prior Research

Prior research does not provide a clear picture of whether performance measurement helps or hinders police effectiveness. Theory behind public sector management reform in the 1980s and 90s suggests performance measurement can improve police performance in delivering service to citizens (O'Flynn, 2007). However, institutional theory suggests that pressure to reflect the values of, and gain legitimacy with, constituents may influence why and how police use performance measurement (Taylor, 2009; Willis et al., 2007). Loose coupling theory suggests performance measurement may actually be counter-
productive, driving staff disengagement and even subversion of policy or the performance that is reported (Brunetto & Farr-Wharton, 2003; Lawton et al., 2000).

The central focus of this thesis is to extend the understanding provided in the literature of the influence of performance measurement on police effectiveness in responding to calls for service. Specifically, I aim to:

Aim 1. examine what police understand effectiveness in responding to calls for service to be;

Aim 2. determine what influence performance measurement has on police effectiveness;

Aim 3. explore why performance measurement has the influence it does on police effectiveness; and

Aim 4. explore the influence of performance measurement on the perspectives and behaviours of police personnel.

To address these aims I employ a mixed methods research design involving both qualitative (Study 1) and quantitative (Study 2) research. In my review of literature on performance measurement (see Chapter 2) and police response to calls for service (see Chapter 3) I found a lack of clarity around the influence of performance measurement on police effectiveness. More information was required before developing hypotheses about the relationship between performance measurement and police effectiveness.
The purpose of Study 1 was therefore to obtain information to assist me to generate testable hypotheses on the relationship between performance measurement and police effectiveness. To obtain this information I conducted 121 semi-structured qualitative interviews with police personnel who deliver or manage the service of responding to calls for service from citizens. In designing the interview schedules I developed specific interviews themes and questions guided by the four aims outlined above. In the interviews I allowed interviewees to provide their own understanding of police effectiveness in responding to calls for service and on the influence of performance measurement. I also prompted interviewees to identify specific examples that illustrate relevant behaviours of police personnel. I utilised findings of Study 1 to generate hypotheses about the relationship between performance information and police effectiveness in responding to calls for service. These hypotheses are:

Hypothesis 1. The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.

Hypothesis 2. Police do not respond to performance measures in a way that influences their organisational legitimacy.

Hypothesis 3. Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.

Hypothesis 4. Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.
I describe these hypotheses in more detail in Section 5.6.

In Study 2 I used time series data collected from New Zealand Police, Queensland Police Service, and New Zealand parliamentary debates to test these four hypotheses. Notably, the time series data allowed me to test the relationship between measures monitored by police and indicators of police effectiveness identified in the literature and by police personnel. In sum, I found that police make only limited use of performance information, and that measuring performance has little influence on police effectiveness in responding to calls for service. However, police personnel maintain a commitment to doing a good job, even when demotivated.

1.4 Thesis Outline

In this chapter I have outlined the key ideas and relationships I explore in this thesis, and described a specific research agenda and methodological approach. In Chapter 2 I provide a more comprehensive justification of the research agenda by way of a literature review on performance measurement in police. In Chapter 3 I build a picture of what has been learned about police response to calls for service from citizens, and of what is required for police to be effective in this aspect of their operation. In Chapter 4 I provide a more detailed description of the methods employed in Study 1. In Chapter 5 I report findings of Study 1 and develop four testable hypotheses about the relationship between performance measurement and police effectiveness in responding to calls for service. In Chapter 6 I describe the methods employed in Study 2. I describe the findings of Study 2 in Chapter 7, addressing the four hypotheses. In Chapter 8 I discuss the key findings of this thesis and their implications for policy and practice.
Chapter 2: Performance Measurement in Police

2.1 Introduction

Since the late 1980s, public police agencies around the world have been striving to improve their performance through various mechanisms including performance measurement, professional standards, training, recruiting, technology, and improving operational methods (Bayley, 1994a, 2008; Fleming & Scott, 2008; Maguire, 2014). Most of these changes concern an ongoing evolution in policing as a result of advances in knowledge and technology. However, as will be discussed in this chapter, the advent of performance measurement challenged the traditional values and nature of police organisations. Performance measurement makes performance transparent to external scrutiny, and using performance information requires different skills and perspectives from those that have traditionally been developed in police leaders.

Many scholars observe that police closely monitor measures related to crime but ignore the many other things that police do (Fleming, 2009; Kelling, 1992; Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; M. H. Moore & Braga, 2003; Skogan & Frydl, 2004). These scholars argue that the absence of broader performance measures limits the effectiveness of police. For example, Maguire (2003), Reiss (1992), and Sparrow et al. (1990) argue that police are less effective than they could be because they have developed a narrow identity as society’s crime fighters, which is reflected in a commensurate narrow focus of the performance measures that police monitor. Fleming (2009) notes that existing performance management tools and measures “are not good at capturing the quality and effectiveness of the work involved and the many contributions that police make to a community’s quality of life” (p. 225).
In this chapter I review prior research that has examined performance measurement in the police context. Some scholars (Cope et al., 1997) identify police reform since the 1980s as a manifestation of public sector management reform. In reviewing public sector management innovations during the twentieth century Van Dooren and Van de Walle (2008) note that only one of these, *New Public Management* (NPM), “included quantitative information in public management on a government wide scale, on an international scale and in all management functions” (p. 21). Given the impact of NPM compared with other performance management innovations, I focus on NPM in particular. In this chapter I start by exploring the historical and theoretical background to NPM before discussing its impact on police. Then, in Chapter 3, I review literature on police response to calls for service from citizens, and identify dimensions of police effectiveness in delivering this service. From these chapters I identify questions that are unanswered in the literature. These questions drive two studies I conducted to determine if police are effective in managing their response to calls for service from citizens.

2.2 The History and Nature of New Public Management

2.2.1 The origin of New Public Management

During the 1980s public sector agencies in western democracies experienced a transformation in management philosophy and structure. This reform movement came to be known as New Public Management (or NPM). The roots of NPM lie in the post-Keynsian welfare state that gave rise to increasingly complex bureaucracies (Mascarenhas, 1993). NPM resulted from public dissatisfaction with the size and associated cost of the public sector, its bureaucracy, and its inability to deliver service (Dawson & Dargie, 1999). This dissatisfaction became particularly evident in the 1970s
(Hood, 1991; Mascarenhas, 1993). The private sector had been performing well in terms of efficiency and service delivery at this time. The stark contrast between private and public sector performance was very evident to both citizens and governments (Beattie & Waterhouse, 2007; Dawson & Dargie, 1999). So, governments adopted private sector principles as best practice (Beattie & Waterhouse, 2007; Cutler & Waine, 2000; Dawson & Dargie, 1999; Dobson, 1982; Hood, 1995; Legrand & Bronitt, 2012; Mascarenhas, 1993).

A convergence of factors triggered the NPM reforms: Public dissatisfaction influenced elections that brought to power conservative governments in Britain (Margaret Thatcher in 1979) and the USA (Ronald Regan in 1980). These governments were philosophically more aligned with private sector models than were their predecessors (Legrand & Bronitt, 2012). However, in Australia and New Zealand more left-wing Labour governments also adopted private sector models (Jesson, 1989). Like Britain and the USA, Australia and New Zealand could not sustain the financial cost of the public sector bureaucracies, and needed to reform the public sector. Growth of multi-national corporations at this time was also driving increasing globalisation of national economies, which further increased the pressure to reform (Hood, 1991).

NPM consisted of three main strands (Butterfield, Edwards, & Woodall, 2004; Cutler & Waine, 2000; Dawson & Dargie, 1999; McLaughlin, Muncie, & Hughes, 2001; Morris & Farrell, 2007). The first strand—streamlining—involves downsizing and outsourcing non-core functions. Sometimes it involved corporatizing and even privatizing whole government departments (de Boer & Evans, 1996; Ferlie, 1992; Morris & Farrell, 2007; Toime & Steele, 1991). The second strand—marketisation—involves establishing
markets. In some instances real markets were created through the outsourcing described above. However, quasi-market models were also established for internal transactions between government departments, and even between different functional groups within the same department (Ferlie, 1992; Mark & Scott, 1992). These markets emulated the private sector to gain the service and efficiency focus evident in the private sector (Dawson & Dargie, 1999; Morris & Farrell, 2007). In the third strand—managerialism—importance was placed on results. Management responsibility was seen as a higher order function than administration (Hoque, Arends, & Alexander, 2004). Public sector managers borrowed from private sector management theory and practices. In particular, they adopted individual and organisational performance measurement frameworks, flexibility in staffing and organisation, and incentives (Dawson & Dargie, 1999; Hoggett, 1996; Morris & Farrell, 2007). Hood (1991) describes these reforms as being based on doctrines of contestability, user choice, transparency, and incentive structures. Such doctrines contrasted sharply with traditional public sector ideas of good administration, which emphasised orderly hierarchies and elimination of duplication or overlap (Hood, 1991).

### 2.2.2 Public choice theory and principal-agent theory

The theoretical basis behind NPM involves public choice theory, principal-agent theory, and competition theory (Hood, 1991). Competition theory (McNulty, 1968) and public choice theory (Buchanan & Tollison, 1984) suggest that governments are unresponsive, inefficient, monopolistic, and unable to reach formal goals. This state occurs because politicians are captured by interest groups and will act in their own self-interest rather than the public interest, and because bureaucrats act in pursuit of self-interest rather than efficiency (Gruening, 2001; Niskanen, 1968). As a result, resources are wasted as
bureaucrats strive to increase their budgets in pursuit of power, status, income, and ease of management (Gruening, 2001; Niskanen, 1968). Public choice advocates seek to persuade governments to adopt policies that support incentive structures based on principal-agent theory to increase efficiency and downsize the state (Mascarenhas, 1993; O'Flynn, 2007).

O'Flynn (2007) describes the central dilemma of principal-agent theorists as being how to get the agent (who is typically an employee or contractor) to act in the interests of the principal (employer) when the agent has an informal advantage over the principal and has different interests from the principal. The challenge for the principal then becomes how to choose an agent and construct incentive structures to align goals in an environment of uncertainty, information asymmetry, a high cost of monitoring, and where incentives exist for agents to shirk. Within NPM, to address this challenge, principals use mechanisms including separating purchasers from suppliers, introducing competition, contracts, performance-based rewards, and performance management and measurement (O'Flynn, 2007). “Performance management and measurement systems developed as a means of monitoring and maintaining organizational control, which is the process of ensuring that an organization pursues action plans that lead to the achievement of overall goals and objectives” (Salem, 2003).

### 2.2.3 Does NPM improve effectiveness?

Pollitt and Bouckaert (2011) summarise research on the effectiveness of NPM as follows:

Elements of the NPM have been widespread, but have they worked? There is no straight-forward ‘yes’ or ‘no’ answer to this, partly because policymakers (and some scholars) start from a strong normative commitment either pro- or anti-
NPM, and they are never likely to agree with each other. However, that is far from being the only reason. It is also the case that it is very difficult systematically to evaluate large-scale public management reforms. (p. 15)

In an early seminal paper on NPM, Hood (1991) observes that “NPM has been most commonly criticized in terms of a claimed contradiction between ‘equity’ and ‘efficiency’ values” (p. 3). Some scholars perceive NPM as eroding “traditional public sector values such as probity, fairness, and equality” (Dawson & Dargie, 1999). Pollitt and Bouckaert (2011) suggest that “Different stakeholders may take very different views of both the justifications and meanings of the reforms, and even of their results” (p. 17). NPM has been influential in effecting change, but has met resistance due to “a collision with the values of the professional cadres” (Lapsley, 2008, p. 80). Dawson and Dargie (1999) argue that resistance from public sector professionals can be explained as a reaction to loss of power. This loss of power occurred when NPM reforms replaced professionals in senior management positions with general managers who were empowered to act in an entrepreneurial manner to deliver outcomes (Lapsley, 2008). NPM reforms were also opposed by unions, whose members were affected by downsizing, cost-cutting, and reduction in job security (Schofield, 2002).

Proponents of NPM, such as Dawson and Dargie (1999), have argued for the economic necessity of NPM and the fairness and efficiency of market models and accountability, citing examples of waste and abuse of power by public sector officials. Examples exist of NPM delivering performance improvements: Toime and Steele (1991) describe improvements in the profitability of New Zealand Post after the New Zealand Government transformed it from a government department into a state owned
corporation. In another New Zealand example de Boer and Evans (1996) describe improvements in telecommunications service and efficiency following reforms that resulted from NPM. Pollitt and Bouckaert (2011) note similar findings in a study of the impact of NPM on British education and health care: “the conclusion was reached that the NPM reforms of the 1980s and 1990s in UK education, health care, and housing has a) raised efficiency, b) improved responsiveness to service users, but c) reduced equity (Boyne et al., 2003)” (p. 18).

Evidence exists that police have failed to fully capture the benefits of NPM reforms (Ashby, Irving, & Longley, 2007). For example, in a study of UK police reform, Savage (2007) notes that, mainly for political reasons, police were shielded from the full effect of NPM reforms. This raises the question: Are police less effective than they might otherwise be because they have failed to fully capture the benefits of NPM reforms? Police did adopt the NPM reform of performance measurement (Bayley, 2008; M. H. Moore & Braga, 2003). However, some scholars argue that the performance measures monitored by police focus too narrowly on measures that reinforce a police identity as society’s crime fighters, rather than reflecting the majority of calls for service citizens make to police (Legrand & Bronitt, 2012; Maguire, 2003; M. H. Moore & Braga, 2003; Reiss, 1992; Sparrow et al., 1990). In this thesis I build on the above study of UK police reform (Savage, 2007) by investigating the influence of performance measurement on police effectiveness in responding to calls for service in Queensland and New Zealand.
2.3 Factors That Influence an Organisation’s Use of Performance Information

Van Dooren and Van de Walle (2008) identify NPM as being pervasive in public management (p. 21). However, they also note that “Performance management needs to be used to have impact” (p. 22). Numerous studies show that managers use performance management and performance information to only a limited extent (Budding, 2004; de Lancer Julnes & Holzer, 2001; van Helden & Ter Bogt, 2001). For example, de Lancer Julnes and Holzer (2001) surveyed over 900 state and government organisations in the USA, and found that even when organisations collected performance measures most did not actually use them to improve decision-making.

Where performance information is used, it can be used for different purposes. Gill (2011) describes performance measurement frameworks as mechanisms that enable members of an organisation to make sense of what is happening in complex organisational contexts. Performance measurement frameworks function as filters that enable sense to be made about what is happening (Weick, 1995). Gill (2011) argues that performance measurement is one way executives can control their organisations. It can influence how members of the organisation perceive the world. For example: “Formal performance reports define events for us that should be significant, such as work targets, and therefore what we need to respond to with action” (p. 407).

Jansen (2008) undertook a series of case studies of Dutch social services agencies. He observed in these case studies a distinction between a citizen perspective on performance, which focuses on citizens in general, and a customer perspective on performance, which
focuses on processes that deliver outputs and outcomes to individuals. Jansen (2008) concluded that:

Politicians have a different perspective on performance, as compared to the internal process perspective and output perspective which are implied in NPM. Managers, especially Production Managers, have an internal perspective, whereas politicians seem to have a citizen perspective and a financial perspective on performance. The citizen perspective is typical of government organisations and it can limit the implementation of more businesslike ways of running such organizations. (p. 190)

Budding (2004) shows that agencies that more strongly embrace NPM reforms are more able to cope with uncertainty in the environments in which they operate. Managers in the organisations studied by Budding (2004) treated planning and control processes as a sports game. This increased their motivation for improving operational management.

Using a mixture of questionnaire and semi-structured interviews, Taylor (2009) studied the use of performance information and performance measurement systems in 40 public sector agencies across Australia. Consistent with the findings of earlier studies (Gregory & Lonti, 2006; Norman, 2002), Taylor (2009) found “that agencies are more likely to use [performance information] for external reporting than for internal improvements” (p. 859). Many respondents viewed the exercise of reporting performance information as a compliance exercise. Furthermore, Taylor (2009) found that performance information used for internal decision-making differed from that reported for external accountability (p. 860). Taylor referred to this as decoupling, which she explained as being a symptom of organisations minimising scrutiny by “maintain[ing] standardized legitimized
structures while their activities vary in response to practical considerations” (Meyer & Rowan, 1977, p. 357).

Studies consistently show that performance measures are more likely to be used within an organisation if they are perceived as valid and relevant to the operational tasks within the organisation (Ammons & Rivenbark, 2008; Askim, 2008; Moynihan & Lavertu, 2012; Radin, 2006; Taylor, 2009). Other studies show that aspects of organisational culture such as openness to innovation, risk-taking, and scrutiny, can increase the use of performance information (Broadnax & Conway, 2001; de Lancer Julnes & Holzer, 2001; Franklin, 2000; Moynihan & Lavertu, 2012) and analytical capability within the organisation (Bourdeaux & Chikoto, 2008; Dull, 2009).

A survey of local government managers in the USA found that public service motivation influences the use of performance information (Moynihan & Pandey, 2010). This study concluded that “performance information use is more likely to be driven by altruism rather than self-interest among government officials” (p. 862). Policing is a public service vocation that exists to serve citizens in need, and police officers are motivated by “a passionate sense of altruism” (D. B. Moore, 1994, p. 198). However, police operate in an environment of intense political interest, because they operate with coercive powers and because of citizens’ fear of crime and citizens’ anger at criminals.

2.4 The Adoption of NPM by Police

The extent and effectiveness of NPM reforms vary between different public sector agencies (Morris & Farrell, 2007). At one extreme entire government bureaucracies, such as those involved in telecommunications (de Boer & Evans, 1996) and transportation
(Edwards, 2009), were privatised. Sectors such as Health and Education were significantly impacted by all three strands—streamlining, marketisation, and managerialism—of the NPM reforms, yet have remained largely public owned (Pollitt & Bouckaert, 2011). At the other extreme agencies such as the police experienced less change, particularly with the streamlining and marketisation strands of NPM (Savage, 2007).

Police unions played a significant role in limiting the impact of NPM management reforms on police. O'Malley and Hutchinson (2007) note that the transfer of business principles to the public police revitalised police unions. These unions resisted NPM reforms that they perceive would disadvantage their members (De Lint, 1998; Finnane, 2002; Kadlec, 2003). For example, in 1992 the British Home Office commissioned a review of the responsibilities and rewards of police (Butterfield et al., 2004). This review released a series of recommendations (the Sheehy report) including a rationalisation of a top-heavy management structure in police, elimination of three ranks, and more flexible pay structures (HMSO, 1993). The Police Federation, Police Superintendents’ Association, and Association of Chief Police Officers immediately initiated a well-orchestrated response (Cope et al., 1997). This response included a large media campaign and a rally of 20,000 police officers at Wembley stadium in July 1993. Under political pressure, in October of the same year, the Home Secretary announced that he no longer accepted significant sections of the Sheehy report (see Cope et al., 1997).

Since the 1990s managerialism changes have become evident in police agencies around the world (Bayley, 1994a; Drake & Simper, 2004; Hoque et al., 2004; Nicholson-Crotty & O'Toole Jr., 2004; Palmer, 1997; Ritsert & Pekar, 2009). Scholars note limitations with
both the extent and the effectiveness of such reforms, and have sought to understand why these limitation occurred. In the remainder of this chapter I review literature on these limitations and their implications for police effectiveness. I begin by discussing relevant theoretical perspectives on NPM in police.

### 2.4.1 Technical-rational theory and institutional theory

Numerous scholars (including Collier, 2001; Crank, 2003; Hoque et al., 2004; Lawton et al., 2000; Vickers & Kouzmin, 2001; Willis et al., 2007) debate the effectiveness of NPM reforms in police. Two theoretical perspectives, technical-rational and institutional, are commonly contrasted in these debates. The technical-rational perspective characterises an organisation by the products and services it delivers (Willis et al., 2007). An organisation’s performance relates to how effectively and efficiently it delivers these products and services. Managers monitor performance measures related to production and delivery to inform decisions that support the effective and efficient delivery of products and services (Hoque et al., 2004; R. W. Scott, 1987). In responding to calls for service, a police organisation may, for example, define its effectiveness in terms of response times, levels of callers’ satisfaction with the police response, or whether offenders are apprehended or convicted due to the work of first-responders. In Study 1 I identify the performance measurement systems used by New Zealand Police and Queensland Police Service and the performance measures monitored within these systems. I also explore how these performance measurement systems influence changes in performance of the organisation.

The institutional perspective (see Meyer & Rowan, 1977) views organisational structures as being influenced by cultural features of their environments. Rather than being driven
to perform according to technical-rational criteria, organisations are judged by how well they respond to beliefs about what they should look like and what they should be doing. Organisations gain legitimacy with their constituents by reflecting these beliefs. This legitimacy improves their prospects for obtaining resources and for their survival (Meyer & Rowan, 1977; R. W. Scott, 2004; Suchman & Edelman, 1996). Organisational legitimacy is achieved when an organisation achieves congruence between the social values it is perceived to demonstrate through its actions and the norms of acceptable behaviour in the societal system in which it exists. “Underlying organizational legitimacy is a process, legitimation, by which an organization seeks approval (or avoidance of sanction) from groups in society” (Kaplan & Ruland, 1991). Institutional goals, such as legitimacy with an organisation’s constituents, may be challenging to measure. In policing such goals may be evidenced by the willingness of citizens to call the police, the behaviour of citizens towards the police, the picture painted in the media about police, how easily a police agency is able to secure the resources it believes it needs to operate effectively, or even the continued survival of the organisation.

Scholars (such as Hoque et al., 2004; Taylor, 2009; Willis et al., 2007) debate whether police use performance measurement frameworks to achieve technical-rational objectives of delivering services more efficiently and effectively, or to achieve institutional objectives such as gaining organisational legitimacy. Willis et al. (2007) suggest that highly institutional organisations, such as police, experience pressures to adopt performance frameworks for structural conformity (isomorphism) with other organisations. DiMaggio and Powell (1983) describe these pressures as taking three forms: (a) mimetic, in which organisations that operate in environments with a lot of uncertainty model themselves on other organisations to enhance their legitimacy and
survive; (b) coercive, where organisations respond to pressure—either formal or informal—from other organisations to adopt specific practices or structure; and (c) normative, where professionals in a field exert their authority as experts, particularly on how members of their profession should work.

Perceptions by an organisation’s employees of misalignment or confusion between technical-rational goals and institutional goals can lead to what is known as loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977). Loose coupling exists where agents are disconnected from organisational goals (Weick, 1990), and can be caused by administrative arrangements or by environmental factors such as the organisation’s social norms (Orton & Weick, 1990). Loose coupling manifests as a gap between formal structures and work practices. Formal rules are often broken, and decisions are often not implemented (P. Smith, 1993). Loose coupling is particularly relevant for organisations with strong institutional characteristics, such as police agencies (Crank, 2003; Klinger & Bridges, 1997; Langworthy, 1986; Lawton et al., 2000; Willis et al., 2007). Lipsky (1980) argues that loose coupling contributes to police being more resistant than other public sector organisations to NPM reforms. Therefore, to understand the effectiveness of performance measurement in police, I present prior research on the relationship between loose coupling and the use of performance measurement in police.

A study of 74 public sector agencies in the United Kingdom concluded that police performance measurement frameworks are counterproductive because they generate “loose coupling which may also lead to a loss of confidence by managers” (Lawton et al., 2000, p. 17). In this study 74 MBA students each selected a different public sector organisation then interviewed members of that organisation and analysed the
organisation’s performance measures using the fieldwork approach of Kanter and Summers (1987). The researchers found that the police agencies studied used externally imposed performance measures that their staff believed were inimical to the needs of the organisation (Lawton et al., 2000). For example, one police officer believed that police headquarters was far removed from the reality of operational policing and was only interested in complying with the demands of external stakeholders (p. 17). One police management team had very little knowledge of its organisation’s performance measures. This team viewed performance reporting as “feel good” (p. 17) exercises. When reporting activities, these activities were simply fitted to match any one of the standards in the framework. Furthermore, anything that might identify poor performance was excluded (p. 17).

A study involving surveying 178 Australian police officers found that lower-ranked police officers “have a reasonably high level of commitment to the Police Service organisation” (Brunetto & Farr-Wharton, 2003, p. 53). The study also found that “police officers have a strong positive regard for their first line managers – namely their sergeants/senior sergeants” (p. 52). However, the study found that first-line supervisors hold negative attitudes towards performance appraisal frameworks. Given the important socialisation role that first-line supervisors perform in an organisation, it may be that first line supervisors condition junior officers in a way that entrenches loose coupling and inhibits the effectiveness of performance measurement frameworks. My research investigates whether this socialisation occurs in New Zealand and Queensland police organisations.
In a mixed methods case-study in Sweden, Andersson and Tengblad (2009) found that NPM reforms face barriers in influencing police activities, because first line supervisors:

“have been socialised into a strong police identity …. they have a clear view of what it means to be a ‘good officer’, but a less clear view of what it means to be a ‘good organizational member’, and act as a representative for the management function.” (p. 48)

Similarly, in a mixed methods study in Australia, Brunetto and Farr-Wharton (2005) also found that police middle managers held negative attitudes following implementation of NPM reforms. Furthermore, these researchers concluded that “If first-level managers perceive a lack of synergy between a written policy and the supporting implementation variables (funding), then it is likely that, to the extent that they have power, they will use it to maintain the status quo” (p. 237). In a Toronto study of approximately 1,000 calls for service from the public, Shearing (1984) also found evidence of police staff breaching policies because of misalignment between those policies and the values of police staff. Although formal policy required police dispatchers to dispatch police units in response to calls for service, dispatchers dispatched units in only 82% of instances. Dispatchers felt they should exercise their professional judgement about whether a unit was really necessary; they also felt motivated to protect units from unnecessary work. This study found that supervisors were often sympathetic to and tolerant of these breaches of policy.

In a review of policing in the USA, Bayley (2008) also found evidence of loose coupling between first-line supervisors and more senior managers. Bayley (2008) characterised police as “a public service bureaucracy with the ethos of blue collar workers. They work in a rule-bound world that discourages initiative and is preoccupied with conditions of
service” (p. 14). Senior managers regard rank-and-file police officers and first-line supervisors “as the source of unhelpful complaints rather than of useful insights” (p. 14). In the resultant organisational culture, senior managers struggle to implement reforms, which are often met with resistance from within police. Bayley (2008) observes that the most significant police reforms were instigated from outside police rather than by police senior managers.

In Study 1 I explore three characteristics of loose coupling that are identified in the literature reviewed in this section on police. These are subversion, which involves acting in ways contrary to the direction given by senior management (Brunetto & Farr-Wharton, 2005), disengagement, which manifests as a lack of alignment with, and buy-in to, the objectives and mechanisms set by senior management (Bayley, 2008; Lawton et al., 2000), and managers tolerating or ignoring breaches of policy by their staff (Shearing, 1984).

The literature reviewed in this section provides evidence of loose coupling in police. This literature also suggests that loose coupling in police may inhibit the implementation of NPM reforms such as performance measurement. However, this does not mean there are no such reforms that have been successful. Police performance measurement reforms have been implemented in numerous countries around the world, including France (de Maillard & Savage, 2012), Britain (Savage, 2007), Portugal (Gomes, Mendes, & Carvalho, 2007), Germany (Ritsert & Pekar, 2009), and Australia (Chilvers & Weatherburn, 2004; Mazerolle, Rombouts, & McBroom, 2007). The police performance measurement reform that is most frequently discussed in the literature is Compstat (Giuliani & Safir, 1998).
2.4.2 Compstat

In 1994, the chief of the New York Police Department, William Bratton, introduced *Compstat*. Compstat is an operational performance review framework that was developed to improve police effectiveness through performance measurement (Bayley, 2008; Giuliani & Safir, 1998). Compstat involves analysts obtaining statistical data pertaining to the city’s police precincts, coupled with police officials holding regular meetings to review the data and discuss police activity and performance. When the initiative first started, analysts presented measures of crime reports, arrests, and police activity at regular weekly and monthly meetings. Chief Bratton would use the data to routinely challenge his precinct commanders about their performance. The action of regularly reporting data combined with routinely challenging police commanders about their performance created an environment with greater accountability for performance than had previously existed within the organisation. Performance statistics radically improved following this innovation. New York City’s mayor, Rudy Giuliani, attributed much of this improvement to Compstat (Giuliani & Safir, 1998; Henry, 2002).

Compstat subsequently spread rapidly throughout the world (with adaptations) as a tool that could be used by police managers to improve the performance of their agencies and reduce crime (Fleming & Scott, 2008; Mazerolle, McBroom, & Rombouts, 2011; Mazerolle et al., 2007; Neyroud, 2008). However, some studies provided explanations other than Compstat for the improvement in crime statistics (Rosenfeld, Fornango, & Baumer, 2005; Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2003). For example, a comparison of homicide trends in New York with homicide trends in the 95 largest U.S. cities suggested that homicide rates were declining across the country anyway (Rosenfeld et al., 2005). “New York’s homicide trend during the 1990’s did not differ
from those of other large cities” (p. 419). Other researchers undertook a survey of all U.S. police agencies with over 100 sworn officers (n=515), plus 100 selected agencies with 50-99 sworn officers, achieving an 86% response rate (Weisburd et al., 2003). These researchers found that strategic problem solving had begun being implemented prior to the introduction of Compstat. Even so, other studies identified examples where operational performance reviews based on Compstat do appear to influence performance measures (Chilvers & Weatherburn, 2004; Mazerolle et al., 2011; Mazerolle et al., 2007; Willis et al., 2007).

Scholars continue to debate police management initiatives such as Compstat – why they work, whether they work, what purposes they serve, what the real motivation behind them might be, what paradigm shift they have brought, and more (Dabney, 2009; Magers, 2004; W. F. Walsh, 2001; W. F. Walsh & Vito, 2004; Willis, 2011; Willis & Mastrofski, 2012). Compstat has been effective in focussing police tactical efforts (Magers, 2004; W. F. Walsh & Vito, 2004), and technical-rational factors have influenced Compstat’s operation (Willis et al., 2007). However, institutional factors also have a strong influence (Willis et al., 2007). In analysing fieldwork data across three U.S. municipal police departments, Willis et al. (2007) concluded that all three departments adopted Compstat to appear progressive and successful. Accordingly, Compstat was implemented in a way that would confer organisational legitimacy and minimise disruption to existing routines. Willis et al. (2007) further concluded that Compstat “was less successful when trying to change those structures and routines widely accepted as being ‘appropriate’” (p. 147).
2.4.3 The evolution of NPM in police around the world

As discussed, Compstat originated in New York in 1994 and, with adaptations, was subsequently adopted in numerous countries (Fleming & Scott, 2008; Neyroud, 2008). However, Compstat is not the only performance measurement framework that has been adopted by police. As I will show in this section, performance measurement in police has evolved and diversified since the 1990s. However, even though the Australian and New Zealand public sectors strongly adopted NPM reforms in the 1980s and 90s, there has been little recent research published on how performance measurement has impacted on and evolved in Australian and New Zealand police agencies. This thesis seeks to address this gap in the literature.

de Maillard and Savage (2012) compared differences between NPM reforms implemented in France and Britain. Britain publishes league tables that score the performance of all constabularies on a range of performance indicators (p. 266). The purpose of publishing league tables is to motivate poorer performing constabularies to improve their performance, thereby increasing their scores. Originally the performance measures focused on indicators such as crime statistics, response times to emergencies, and crime detection (clear-up/resolution) rates. However, in the late 1990s the Association of Chief Police Officers (ACPO) argued for a broader range of measures (p. 368). The Audit Commission responded by extending the suite of indicators to include measures of customer satisfaction (Savage, 2007). Since the late 1990s performance measurement in UK police has continued to grow and broaden (Legrand & Bronitt, 2012). In 2010, influenced by the U.S. reforms discussed by Bayley (2008), the Home Office

\[\text{1 After the study by de Maillard and Savage (2012), Britain made changes to their practice of publishing league tables.}\]
created new roles of Police and Crime Commissioners. These elected civilian roles were given the power to set policing priorities for their local constabularies. They were also charged with holding police to account for their performance (Home-Office, 2006).

Similarly, the reforms undertaken by French police focussed on accountability for performance. de Maillard and Savage (2012) describe both British and French police as having experienced managerialism reforms in response to political agendas that embraced both the ethos of NPM and public discourse of being “tough on crime” (p. 374). However, rather than publishing league tables of performance indicators, France instigated broader public sectors reforms that impacted police organisations. de Maillard and Savage (2012) observed that police in France introduced elements of NPM in the early 1990s, such as performance indicators with targets. However, momentum slowed during the 1990s.

More recently, two particular developments in France, instigated external to police organisations, triggered more extensive adoption of managerialism reform. The first of these was a change to the framework within which public spending is organised. The loi d’orientation relative aux lois de finances, Organic law on finances laws (LOLF) was enacted in 2001, and came into effect in 2006 (de Maillard & Savage, 2012, p. 373). This moved public sector agencies away from line-by-line budgeting to a framework where programme managers and ministers are required to make commitments to achieving specific objectives and targets. More recently, managerialism has been invigorated by a further reform known as Révision générale des politiques publiques. This reform involves mixed teams of ministerial inspectorates and consultants auditing and reviewing public sector agencies. For the national police and the national gendarmerie this review produced
recommendations consistent with NPM innovations, including “the further civilianisation of the police and the outsourcing of non-essential responsibilities” (p. 373).

These reforms met with resistance. Police unions and police commissaires in France criticised a NPM-style focus on results (de Maillard & Savage, 2012, p. 376). Furthermore, evidence was found of French police officers acting in ways that artificially improve measures of crime and clear-up rate (Matelly & Mouhanna, 2007). Also, France lacks the transparency of Britain’s league tables, because performance indicators are only published at national level with no data available on sub-national results (de Maillard & Savage, 2012, p. 377).

Bayley (2008) argues that police in the USA have become more open to external scrutiny compared to other countries. Just as de Maillard and Savage (2012) observes that externally instigated reforms have impacted French police, Bayley (2008) also observes that most significant reforms affecting U.S. police have been externally driven. He postulates that far-sighted police leaders sought to drive improvements in police effectiveness by creating openness to criticism of the police establishment (Bayley, 2008, p. 15). U.S. police have harnessed science in the form of empirical evaluation, are supporting study of police by people not employed by police, are collaborating with outsiders, and are implanting evaluation within police departments (Skogan & Frydl, 2003).

Gomes et al. (2007) observed that management and measurement of performance within law enforcement agencies in Portugal have not developed to the same extent as in other public sector organisations. They argue that factors such as ambiguity of objectives and the political context in which law enforcement operates make comparability of
performance information difficult. Furthermore, goals in policing are often complex and can be difficult to measure. Such factors pose challenges for performance measurement in police (Gomes & Mendes, 2013).

Ritsert and Pekar (2009) note that reforms in Germany occurred later than in western countries of Anglo origin. NPM reforms have only been partially adopted by German police, who take a pragmatic, rather than philosophical, approach to NPM. Ritsert and Pekar (2009) describe an approach of using a “NPM-tool-kit” (p. 40) from which police are selective about which tools they apply. As a result, German police have a hybrid organisational culture and structure. This is a traditional administrative culture, but with elements of NPM.

Coleman (2008) suggests that performance measurement in police has limited impact because police managers lack the necessary skills to use performance information effectively. In a study of strategic performance measurement across 39 Canadian police organisations, Coleman (2008) found “inadequate knowledge transfer to, and education of, police leaders with respect to the value of strategic management and the measurement of organizational performance that is necessary to improve organizational performance” (p. 319). This inadequate knowledge transfer to police leaders may be because they have been so socialised into thinking like police officers that they are unable to think like managers (Andersson & Tengblad, 2009). Because police management consists mainly of promoted police officers rather than managers with broader experience, police may be less effective than other organisations in using performance information to manage and improve their performance. Whereas this argument may be plausible it is far from conclusive, because the study by Coleman (2008) is a single study involving a survey of
only 39 respondents, all from Canada. In my research I test and extend the findings of Coleman (2008) by investigating the use of performance information by police managers in Australia and New Zealand.

The diversity and limited adoption by police of NPM reforms discussed in this section suggest it is particularly challenging to reform police in this way. Performance frameworks can be used, and possibly abused, to make a police agency appear legitimate rather than to help police improve the effectiveness of their service delivery. Performance measurement frameworks as implemented by police may be ineffective or even counterproductive. Given these concerns, it is worth considering whether police should measure their performance at all. This will be examined in the next section.

2.5 Why Measure Performance at all?

The literature reviewed in Section 2.4 suggests that measuring performance in police may increase the degree of loose coupling between senior management, middle managers, and staff. This loose coupling limits the ability of senior management to control the organisation. Additionally, measuring performance creates a risk of attracting negative media coverage when performance indicators reveal poor performance (Taylor, 2009). These risks create disincentives for senior managers to want to measure performance.

M. H. Moore and Braga (2003) suggest three reasons why police executives would want to measure performance. First, external stakeholders, such as citizens, governments, and persistent interest groups, will demand an account of what police have done with the resources entrusted to them (P. Smith, 1995). Second, it is the ethically and morally right thing to do (M. H. Moore & Braga, 2003). Executive members have been entrusted with
the responsibility for deploying the police resource on behalf of their stakeholders. Third, police executives need to embrace external accountability to leverage it to drive the organisation to high levels of performance (Bayley, 2008; M. H. Moore & Braga, 2003).

The early theoretical basis for NPM viewed performance measurement as a mechanism to make bureaucrats efficient and accountable by monitoring performance against pre-set targets (Bevan & Hood, 2006; Newberry & Pallot, 2004). More recent research suggests that performance measurement can enhance performance irrespective of the performance targets being set: In a recent survey of 101 public sector organisations Speklé and Verbeeten (2014) found that organisations that make exploratory use of performance information tend to enhance performance. Performance is enhanced independently of the whether performance targets are set (p. 140).

Studies of the impact of NPM reforms on police in the Australian police context found that they serve the purposes of encouraging efficiency of resources and legitimising the police to citizens (Fleming & Lafferty, 2000; Hoque et al., 2004). Instances exist of police implementing performance measurement in reaction to, or as a preventative measure to avoid crises of police organisational legitimacy (E. Davis, 2002; Fleming & Lafferty, 2000; Sarre & Prenzler, 2005), or failures of business processes (Corboy et al., 2005; Jones, 1995). Measuring performance demonstrates accountability, and may also alert police managers to emerging risks that, if ignored, may turn into crises of organisational legitimacy, such as occurred in the cases discussed in Chapter 1 of Iraena Asher and Cherie Cundy (Corboy et al., 2005; Hutton, 2012; Leask, 2012).
2.6 Conclusion

The literature reviewed in this chapter raises a number of questions about the impact of performance measurement on police effectiveness. Benefits exist in police measuring their performance (Hoque et al., 2004; M. H. Moore & Braga, 2003), but limitations (Coleman, 2008; Legrand & Bronitt, 2012) and risks (Taylor, 2009) exist as well. Police agencies are strongly institutionalised (Crank, 2003). As a consequence, it may be that police implement performance measurement frameworks to gain organisational legitimacy by appearing progressive, rather than to improve service delivery by improving their management of processes and resources (DiMaggio & Powell, 1983; Taylor, 2009; Willis et al., 2007).

Moynihan and Pandey (2010) argue “that understanding public employee use of performance information is perhaps the most pressing challenge for scholarship on performance management” (p. 849) because of the effort governments have given to creating performance data to improve governance, and because there is much we do not know about the factors associated with the use of that information. Van Dooren, Bouckaert, and Halligan (2010) suggest that barriers to using performance information exist at three levels: (a) cultural barriers within an organisation, (b) institutional barriers created by frameworks used by professions which influence how choice is framed and sense-making is shaped, and (c) individual psychological barriers such as limited cognitive ability. Police are highly institutionalised (Crank, 2003) loosely coupled (Willis et al., 2007) organisations whose management is dominated by members of a single profession (Andersson & Tengblad, 2009; Jacobs, Zettlemoyer, & Houston, 2014). Therefore at least the first two of the three barriers suggested by Van Dooren et al. (2010) are relevant to police organisations. Individual psychological barriers may also exist.
Coleman (2008) suggests that police managers may lack the skills to use performance information effectively. This lack of skills may result from police leaders being so socialised into thinking like police officers that they are unable to think like managers (Andersson & Tengblad, 2009). However, evidence supporting these ideas is limited.

In order to test whether police are effective in their use of performance information it is necessary to identify a specific context about which criteria for police effectiveness can be defined. In the next chapter I review literature on a specific service provided by police: responding to calls for service from citizens. By understanding the key ingredients that make police effective in responding to calls for service, I am then able to identify specific criteria for police effectiveness. I use these criteria to frame specific research questions for my studies.
Chapter 3: Police Response to Calls for Service

3.1 Introduction

Since the 1980s, particularly in democratic countries, significant change and innovation has occurred in public policing (Bayley, 1994a; Maguire, 2014; Weisburd & Braga, 2006). Technological innovation gave new tools to police, such as computer systems, DNA testing, CCTV, and enhancements in telecommunications (Chan, 2001; Reeder, 1999). Also, police learned more about what works in combating crime (Telep & Weisburd, 2012). This led to new proactive approaches to policing, such as problem-oriented policing (Eck & Spelman, 1987; Goldstein, 1990), intelligence-led policing (Ratcliffe, 2003), and community-oriented policing (Dammer et al., 2006; Goldstein, 1987b; Greene & Mastrofski, 1988). Furthermore, like other public sector agencies, since the 1980s police have been implementing NPM reforms aimed at making police more effective, efficient, and accountable (Bayley, 1994a; Coleman, 2008; Cope et al., 1997).

Most of the recent innovations in policing aim to make police more proactive. However, citizens expect police to provide a reactive service as well (Sparrow et al., 1990). Police are regularly called to attend many different types of situations. These include natural disasters, traffic accidents, crimes, public disorder, and numerous different circumstances where someone thinks that a competent authority figure is required (Bayley, 1990, 1994b; Sparrow et al., 1990; Terrill, 2003). If people fear that police will not respond when and how required, they will become dissatisfied and less willing to assist police (Eckblom & Heal, 1982; Skogan, 2006; Sparrow et al., 1990).
Through the second half of the twentieth century police agencies made numerous reforms to improve their response service (Bayley, 1994a; Maguire, 2014; Weisburd & Braga, 2006). Mechanisms employed to improve service include the creation of centralised specialist centres for receiving calls and dispatching police units, with recognised nationally standardised phone numbers for emergency and for non-emergency calls (Manning, 1992; Mazerolle et al., 2002; Reiss, 1992; Sparrow et al., 1990), the introduction of performance measures and standards (Brown & Coulter, 1983; Furstenberg & Wellford, 1973), and the use of mathematical modelling and planning of resource deployments (Chen & Henderson, 2001; Green & Kolesar, 1989; Larson, 1972b).

In this chapter I explore the key ingredients that make police effective in responding to calls for service. I begin by discussing what constitutes police response to calls for service, and why it is important for police to be effective in this area of their operation. I show that there are many dimensions to effectiveness. To determine whether police are effective in these different dimensions I consider prior research examining the relationship between police actions and police effectiveness. This research reveals that police face many and complex challenges to be effective in responding to calls for service, raising questions about the influence of the performance measurement frameworks that police use to manage this complexity. I conclude the chapter by identifying these questions and proposing a research design to answer them.
3.2 The Challenges Police Face in Responding Effectively to Calls for Service

Ensuring appropriate response to public calls for service is a constant challenge for police around the world (Corboy et al., 2005; Gunal, Ongo, & Pidd, 2007; Kuhn & Hoey, 1987; Moskos, 2007; Sacks, Larson, & Schaack, 1993; Sparrow et al., 1990). The volume of calls to police can vary markedly throughout the day, making it difficult for police to answer all calls in a timely manner (Gunal et al., 2007). Calls for service vary in their urgency, importance, and relevance to police (Moskos, 2007). Police must often prioritise which incidents to attend, because there may be more incidents requiring police attendance than there are police units available (Moskos, 2007). Decisions about the priority of a given incident are often based on limited information, and there may be potentially serious consequences if the wrong decision is made (Moskos, 2007; Sampson, 2004). Even if a unit is available to attend the incident, it may be located at a distance from the address it has been dispatched to, or traffic may delay its arrival. Upon arriving, the circumstances encountered may differ from what was expected (Klinger & Bridges, 1997).

Many calls to police are not legitimate, either because they are hoaxes, wrong numbers, or accidentally made by children playing with the telephone (Brumfield, 2014; Sampson, 2004). The nature or urgency of legitimate calls can be difficult to assess, and the assessment made by the police call-taker may differ from that of the caller (Moskos, 2007). Furthermore, police decisions—both those of the call-taker and the officers attending the scene—will not suit everybody, and some people may feel that the police’s decisions were discriminatory (Mladenka & Hill, 1978; Weitzer & Tuch, 1999).
The challenges illustrated above show that complex and interrelated components comprise the service of responding to calls for service from citizens. The following discussion examines each of these components. The discussion begins with an examination of the reasons why citizens report to police, followed by an analysis of management of calls by police and dispatch of units to the scene. Lastly, a review of previous research that has examined those factors is undertaken.

### 3.2.1 Citizens reporting to police

**Why victims report crimes to police**

Society often perceives the police as the gate-keepers of the criminal justice system (Buzawa, Austin, Bannon, & Jackson, 1992; Frisch & Caruso, 1996). However, Hindelang and Gottfredson (1976) argue that victims serve as the gatekeepers. By deciding whether to report crimes to the police, victims control mobilisation of the criminal justice system, of which police are a part. Scholars observe that police are frequently called back to the same location or victim, and that the overall volume of crime and/or calls for service could be reduced if repeat victimisation could be prevented through an effective initial response (Farrell & Pease, 1993; Pease, 1998; Townsley, Homel, & Chaseling, 2000). Gottfredson and Gottfredson (1980) and Hindelang and Gottfredson (1976) argue that, because victims are the first major decision makers in the process of calls to police for service, it is important to understand the factors that influence which victims report to police. In this section I discuss three factors identified in the literature that potentially influence victims to report crime to police: seriousness of the offence, numbers of police, and the previous experience of the victim.
Studies consistently show that the most important determinant of whether an offence is reported to the police is the seriousness of what occurred, in particular, the resultant physical injury, property loss, or emotional impact (Carach, 1997; Felson, Messner, & Hoskin, 1999; Goudriaan, Wittebrood, & Nieuwbeerta, 2006; Hart & Rennison, 2003; Kääriäinen & Sirén, 2011; MacDonald, 2001; Skogan, 1984, 1994a; Van Kesteren, Mayhew, & Nieuwbeerta, 2001). Characteristics of victims also influence whether victims choose to report the offence, but this has less influence than offence seriousness after controlling for the nature of the offence and characteristics of the local community (Cao, Frank, & Cullen, 1996). The most common reason victims give for not reporting to the police is that the matter is too trivial (Morrison, Smith, & Gregg, 2010). Victims also frequently say they did not report the offence because the police either could not or would not have done anything. This is often because they perceive police as being too busy (Morrison et al., 2010). Victims also view many instances of offending as a private matter, and do not see police intervention as appropriate (Felson, Messner, Hoskin, & Deane, 2002; Mayhew & Reilly, 2007; Morrison et al., 2010; Van Kesteren et al., 2001).

Some scholars believe that by increasing police numbers recorded crime will increase, due either to increased police capacity to patrol and detect crime, or increased willingness of citizens to report crime if they believe police have the resources to respond (Bayley, 1985). Others theorise that increasing the number of police officers will deter offending, resulting in an associated reduction in recorded crime (Marvell & Moody, 1996; Walker, 1992). Prior research examined the influence of the size of a police force on levels of recorded crime and apprehensions of offenders (Benson, Rasmussen, & Kim, 1998; Chamlin & Langworthy, 1996; Drake & Simper, 2000, 2005; Robinson, Young, & Cameron, 1989). These studies show apparently contradictory results, possibly because
their methodologies are not strong designs. Weisburd and Eck (2004) argue that such studies generally do not control for the factors that influence the setting of police strength. Some scholars argue that the different ways police agencies deploy their staff may influence crime volumes (Benson et al., 1998; P. J. Cook, 1979; Sherman, 1997). However, overall research to-date provides little support for the view that simply increasing total police numbers has a significant impact on the volume of recorded crimes (Cameron, 1988; Eck & Maguire, 2005; Levitt, 1998; Robinson et al., 1989).

So far in this section I have discussed the influence of offence seriousness and police strength on victims deciding to report crime to police. Victims’ previous experiences can also influence their decision to report crime to police. Thibaut and Walker (1975) identify a human desire to maintain significant control over major life decisions. It follows that a person’s behaviour will be influenced by their perception of whether a given behaviour will allow them to maintain control over their life. Social learning theory predicts that a person’s behaviour is conditioned by previous experience of the consequences of a given behaviour (Akers, 1973). Social learning theory, first described by (Bandura & Walters, 1963), posits that learning is a cognitive process in which people learn behaviours, not only by incentives that reinforce behaviours with favourable consequences, but also by observing consequences others receive for behaviours.

Based on these ideas of Akers (1973) and Thibaut and Walker (1975), Hickman and Simpson (2003) hypothesised that when police act according to the preferences of victims, those victims will be more satisfied with police and more likely to report offences to police. To test this hypothesis, Hickman and Simpson (2003) analysed data from the Metro-Dade County (Florida) experiment on domestic violence (A. M. Pate & Hamilton,
1992; A. M. Pate, Hamilton, & Annan, 1991, 1994) and found that when police decisions reflect the victim’s preferences the victim is more likely to report to police again in the future. It follows that, if victims reporting to police assists the police to be effective, then the extent to which police are influenced by the preferences of victims will be proportional to police effectiveness. I therefore explore whether victims reporting to police does assist police to be effective.

**What gets reported to police**

Although citizens often report crime or suspicious activity that could be crime-related, most calls for service to police are not crime-related (Cumming, Cumming, & Edell, 1965; D. A. Smith & Klein, 1984; Sparrow et al., 1990; Walker, 1992). Police workload from calls for service is dominated by other activities (Criminal Justice Commission, 1996; Hoffman, 1971; Walker, 1992). Such activities include attending natural disasters, vehicle crashes and other emergencies, searching for missing people and rescuing people in danger, dealing with drunk people and people with mental health issues, attending sudden deaths, and restoring order in disputes (Bayley, 1990, 1994b; Sparrow et al., 1990; Terrill, 2003).

**3.2.2 Handling calls from the public**

During the latter part of the twentieth century, in an effort to cope with an increasing volume of emergency calls for service, police agencies around the world progressively implemented centralised call-handling centres (Manning, 1992; Reiss, 1992; Sampson, 2004). Emergency calls that were once received at local police stations are now typically received at centralised police call centres. These centres receive calls and dispatch police using computer-aided dispatch systems (Manning, 1992; Reiss, 1992; Sampson, 2004).
Around the end of the twentieth century, to make it more convenient for citizens to make non-emergency calls to police, and to prevent non-emergency calls from congesting emergency phone lines, police agencies began creating single non-emergency phone numbers and centralised non-emergency call centres (Bayley, 1998; Mazerolle et al., 2002). Centralised call centres create economies of scale that enable operations research algorithms to be applied to predict call volumes and set staffing levels. This assists police to effectively manage specific service levels related to call-handling delays (Chen & Henderson, 2001; Kuhn & Hoey, 1987; Larson, 1972b; Mabert, 1985).

As evidenced at the start of this thesis, through the failure of police to prevent the deaths of Iraena Asher (Leask, 2012) and Cherie Cundy (Hutton, 2012), police effectiveness in handling calls for service involves more than just coping with the volume of calls police receive. In reviewing data from police systems in Baltimore, Maryland, for example, Moskos (2007) identified numerous challenges police face in handling and quantifying calls for service. Some calls are made accidentally, particularly from mobile phones or by small children playing with the telephone. In the case of intentionally made calls, citizens can provide misleading information. Citizens may call police as a hoax or may exaggerate the seriousness of events to prompt a faster police response. Alternately, they may be ignorant about whether the matter is civil, criminal, or private. Police may receive multiple reports of the same event, and may not always immediately recognise that two calls relate to the same event. Moskos (2007) also found that reports deemed illegitimate by police are less likely to be recorded. In an earlier survey on police patrol and activity Whitaker (1982) found that officers in the field often do not record matters they become aware of.
These factors affect what police record about the calls for service they receive. It follows that police may struggle to measure their effectiveness in responding to calls for service because calls for service can be hard to measure. Studies consistently identify difficulties in accurately classifying and quantifying calls for service (Bursik & Grasmick, 1993; Gilsinan, 1989; Klinger & Bridges, 1997; Moskos, 2007; Reiss, 1971; Sampson, 2004; Whitaker, 1979). For example, Klinger & Bridges (1997) identify multiple ways in which matters become known to police, and argue that this makes it difficult to measure calls for service in a consistent way.

### 3.2.3 Dispatch

Upon receiving a call for service at a centralised police call-centre the police call-taker logs a job into a computerised dispatch management system (Chen & Henderson, 2001; Guidi, Townsley, & Homel, 1997). The police call-taker classifies the job type, and assigns a priority for dispatch. A dispatcher, who may or may not be the same person as the call-taker, then attempts to locate and assign a suitable unit to attend the job (Gunal et al., 2007; Moskos, 2007). Sometimes police agencies establish specialist teams to respond to certain types of jobs. Such teams may, for example, include *special weapons assault teams* (SWAT), also known as *armed offenders squads* (AOS), or specialist burglary, fraud, or sexual assault squads/teams. However, most often, a general duties/patrol unit is initially dispatched to attend (Reiss, 1992). To varying extents different police agencies operate *split-force* strategies that partition staff responding to calls for service from those undertaking proactive tasks (Mazerolle et al., 2002; Tien, Simon, & Larson, 1979).
Many police agencies measure their timeliness of response to dispatched jobs. Early literature assumed that rapid response was a significant factor in the apprehension of suspects and in the satisfaction of citizens (Farmer, 1984; Larson, 1972a; Manning, 1977). Some studies of police data (Clawson & Chang, 1977; Isaacs, 1967) supported this view by showing that rapid response to crimes significantly increased the probability of apprehending suspects.

Studies from the late 1970s and 1980s, however, showed that if police do not arrive within the first minute of a crime occurring rapid police response has negligible impact on the likelihood of apprehending suspects (G. W. Cordner, Greene, & Bynum, 1983; Reiss, 1992). In particular, two seminal studies (Kansas City Missouri Police Department, 1977; Spelman & Brown, 1984) showed that rapid response to reports of crimes that were still in progress resulted in apprehension of offenders in about one third of cases. However, these studies also found that most offences are reported to police well after the offender has departed. In such instances rapid police response makes little difference to the likelihood of apprehending offenders. In a more recent study Cihan, Zhang, and Hoover (2012) analysed data from the Houston Police Department on calls for service and found that rapid response increases the probability of apprehension in instances where the burglary is still in progress. However, because most offences are reported to police well after the offender has departed, Telep and Weisburd (2012) in reviewing research to-date conclude that, although rapid response can lead to apprehending suspects, there is no evidence that rapid response to most calls increases apprehension rates or reduces crime.

Even though there is usually a delay between a crime occurring and being reported to police, citizens like to believe that the police will arrive quickly if they are called
Citizens’ perceptions of the speed of police response to calls for service have been shown to be a strong predictor of satisfaction with police performance (Parks, 1984; T. Pate, 1976; Percy, 1980). However, these and other studies (Brown & Coulter, 1983; McEwen, Connors, & Cohen, 1986; Tien et al., 1979; Worden, 1993) also found that citizens’ satisfaction with police service is less influenced by the actual speed of police response than it is by whether the response matched the expectations of the caller. More generally, in a survey of citizens who had recently had contact with police, Reisig and Chandek (2001) found that disparity between citizens’ expectations and the service they receive from police is inversely related to satisfaction with the encounter. This result held for both citizens who voluntarily contacted police to report offences (n=211) and those who were stopped by police and given traffic citations (n=397). Scott (1989) suggests that citizens’ expectations are influenced by what police tell people to expect. However, literature is lacking on evidence about the extent to which police can or do manage citizens’ expectations about speed of police response.

Some evidence does exist that recent contacts with police influences the perceptions of citizens in respect to how quickly police respond. For example, Parks (1984), in a study involving interviews with citizens in 44 residential neighbourhoods served by 29 separate police agencies, found that a prior unsatisfactory police contact has a large negative impact on citizens’ perceptions of response. More recently, Skogan (2006) and Myhill and Bradford (2011) found a similar relationship and that it is asymmetrical: Unsatisfactory prior encounters with police strongly influence perceptions, but satisfactory encounters do not.
The knowledge that meeting citizens’ expectations of police response timeliness has more influence on satisfying citizens than does actual timeliness of police response, enabled police to develop strategies such as differential police response (DPR). DPR strategies differentiate between calls for service in a systematic way. They involve providing a wider range of response options than the traditional approach of immediately dispatching a general-duties unit to attend as quickly as possible. For example, police call-takers may arrange with the caller to attend at a later time, or for civilian personnel, rather than police officers, to attend certain types of jobs (Worden, 1993). For certain types of jobs the call-taker may request that the caller comes to a police station, or may decide no action is required (Moskos, 2007; Sparrow et al., 1990).

Because DPR strategies differentiate between calls for service, they create a real or perceived opportunity for police to discriminate against certain types of citizen (Worden, 1993). However, studies consistently demonstrate that race or the socioeconomic status of citizens or neighbourhoods have little influence on police decisions (Mladenka & Hill, 1978). In fact, DPR strategies appear to produce greater equity by systematising the criteria police use for dispatch decisions (McEwen et al., 1986; Worden, 1993). Furthermore, DPR does not appear to reduce citizens’ satisfaction with police service (Brown & Coulter, 1983; McEwen et al., 1986; Mladenka & Hill, 1978; Thurow, 1970; Worden, 1993).

The adoption of DPR strategies have been influenced by a number of key drivers. These include the need to cope with increasing volumes of calls for service (Bayley, 1998; Gay, Schell, & Schack, 1977; Sparrow et al., 1990; Worden, 1993), fiscal pressures to make cutbacks (Levine, 1985; Sacks et al., 1993), and NPM reforms (see Chapter 2) to make
the public sector in general, and police in particular, more effective (Eck & Spelman, 1987). As with the innovations discussed above regarding centralised call centres and nationally standardised phone numbers, the adoption of DPR strategies demonstrates that police do innovate to improve their efficiency and effectiveness. However, these innovations create other challenges for police to manage, such as increased demand and more complicated dispatch options. Also, police face challenges and opportunities when they arrive at the scene of incidents to which they are called. This is the focus of the next section.

3.2.4 Attendance at the scene of the call out

The importance of first-responders to resolving crimes

A number of studies in the late 1970s and early 1980s (Eck, 1983; Greenwood, Chaiken, Petersilia, & Prusoff, 1975; Greenwood, Petersilia, & Chaiken, 1977; Skogan & Antunes, 1979) suggested that police investigations have little impact on either the volume of crime or the apprehension of offenders. However, the quality of work in gathering evidence by police first-responders to a crime scene has significant influence on the likelihood of convicting suspects who are apprehended (Forst, 1982). Studies such as Eck (1983), Greenwood et al. (1975), and Greenwood et al. (1977) consistently show that follow-up investigations are unlikely to be successful unless the officers who first respond obtain information about suspects from citizens.

Procedural justice

As well as influencing the likelihood of convicting suspects, the behaviour of police first-responders can also influence citizens’ satisfaction with police and citizens’ obedience to the law. During the late 1970s and early 1980s, studies such as that by Tyler and Folger
(1980) found that citizens’ attitudes towards police were influenced favourably when officers created a perception that they had dealt with the matter properly and fairly. Police officers created this perception by appearing to act in an objective manner and by spending time examining the circumstances of the encounter, questioning suspects, recovering stolen property, and/or completing a formal report, even when the outcome was not favourable for the citizen (Tyler & Folger, 1980). Conversely, citizens developed unfavourable attitudes when they perceived a lack of fairness or when officers simply asked a few questions, then departed without taking further action (Parks, 1984; Tyler & Folger, 1980). These studies suggest that procedural justice theory is applicable to policing.

Thibaut & Walker (1975) had earlier described the concept of procedural justice in the context of resolving disputes. Thibaut & Walker (1975) posited that the procedures used to resolve a dispute have an impact on satisfaction that is independent of the outcomes received. To test this hypothesis in a policing context, Tyler and Folger (1980) conducted 184 telephone interviews from a randomised sample of residents living in Evanston, Illinois. Sixty six percent of respondents stated they had called police at least once in the previous five years to report a crime, make a complaint, or request assistance. These interviews enquired both about the outcome of the contact with police and how fairly police had responded. The study found that callers developed a level of satisfaction with police that related to their perceptions of how fairly they were treated, irrespective of whether police solved the problem to which they were called or sanctioned the citizen in some way for violating the law (Tyler & Folger, 1980). In a later experiment conducted in Milwaukee in 1987 and 1988, to determine whether either arrest or non-arrest deterred re-offending, cases of misdemeanor domestic assault where probable cause to arrest
existed were randomly assigned to police action of arrest or non-arrest (Paternoster, Braume, Bachman, & Sherman, 1997). Consistent with the findings of Tyler and Folger (1980), in analysing data from 479 interviews of persons arrested in the Milwaukee experiment, researchers found a reduction in repeat domestic assault among arrestees who perceived that police had treated them respectfully. Taking the time to listen to the offender’s version of a story had a powerful effect on the likelihood that the offender would reoffend (Paternoster et al., 1997).

Tyler and his associates subsequently showed that police behaviour influences police legitimacy which, in turn, influences citizens’ obedience to the law (Sherman, 1997; Tyler, 1990, 1994; Tyler & Huo, 2002). This is an important finding, both empirically and theoretically, because it links citizens’ subjective attitudes and perceptions with objective crime prevention outcomes. Although early research on procedural justice in police was US-based, more recent studies in other countries, including Australia, Britain, and India, replicated similar results (Hinds, 2009; Hinds & Murphy, 2007; Hough, Jackson, Bradford, Myhill, & Quinton, 2010; Jonathan-Zamir & Weisburd, 2011; Mazerolle, Antrobus, Bennett, & Tyler, 2013; Mazerolle, Bennett, Davis, Sargeant, & Manning, 2013; Tyler & Murphy, 2011). Replicating results in different contexts and locations generalised the early U.S. findings. These studies reinforce the importance of procedural justice for perceptions of police effectiveness in responding to calls for service.

**The influence of political pressures**

As discussed, police can be effective in technical-rational terms by apprehending offenders as a result of thoroughness in gathering evidence, and by preventing reoffending
through demonstrating procedural justice. However, police face institutional pressures that arise from the politicised context in which they operate. Governments seek to maintain their legitimacy with citizens by placing pressure on police to reduce crime statistics and satisfy citizens (Seidman & Couzens, 1974). For example, a survey of Kentucky police chiefs (n=115) found that political pressure influenced police operational decisions (Tunnell & Gaines, 1992), and concluded that “Kentucky police chiefs must cope with varying levels of political pressures and interference, some legitimate and some illegitimate” (p. 14).

Political movements can influence government policies and police practice. For example, Tierney (1982) observed that “Wife beating has become the object of media attention and government policy, not because of an increase in its frequency, or because the public has become more concerned, but because a social movement developed in the 1970s to help battered women” (p. 207). Felson (2005) found that police can be prevented from responding to calls for service in the most effective way partly because of the influence of this social movement. He cites a study in the early 1980s involving Minneapolis Police Department (Sherman & Berk, 1984) that examined the effects of different actions taken by police first-responders to domestic assault, concluding that aggressive arrest practices reduced the likelihood of repeat offending. Felson (2005) observes that these findings “appealed both to conservatives, who wanted to get tough on violent criminals, and to feminist activists, who wanted to get tough on violent husbands” (p. 564). In the resultant political climate many police agencies, particularly in the USA, instigated a policy of mandatory arrest in instances where prima-facie domestic assault could be established. Many U.S. states even legislated to remove police discretion not to arrest (Backman &
Because the results of the Minneapolis study aligned with political pressures, these results influenced public policy. However, where evidence does not align with political pressures, it appears to have less influence. Subsequent studies failed to replicate the Minneapolis (Sherman & Berk, 1984) study results (Berk, Campbell, Klap, & Western, 1992; Felson et al., 2005; Sherman, 1992). For example, Felson, et al. (2005) and Langan and Innes (1986) separately analysed longitudinal data from the U.S. National Crime Victimisation Survey, which interview respondents six times over a three-year period. Both studies found that reporting domestic violence to police significantly reduces the likelihood of repeat offending. In the most recent of these studies, Felson et al. (2005) included 2,564 respondents who were victimised by their spouses. This study concluded that “reporting has a fairly strong deterrent effect, whereas the effect of arrest is small and statistically insignificant” (p.563). More recently, Eve Buzawa and colleagues reviewed the consequences of mandatory arrest policies for domestic violence. They found “negative consequences, including the increase of dual arrests where both parties are arrested, and victim dissatisfaction coupled with failure to report when their preferences are not followed.”(Buzawa, 2012, p.83). Despite such results, mandatory arrest policies and legislation generally remain in place. Police officers are required to arrest, even in situations where such action conflicts with both their judgement and the wishes of the victim.

So far in this chapter I have reviewed literature on the complexity of the task of responding to calls for service, the influence of police behaviour on citizens, and influence
of political pressures. Adding to the complexity police must contend with, technological innovations provide both opportunities and challenges for how police operate in delivering a service that responds to calls for service from the public. In the following section I explore how police have responded to such technological innovations.

### 3.2.5 The impact of technological innovation

In an attempt to increase efficiency and effectiveness, police have, to varying extents, embraced technological innovations (Chan, 2001; Maguire, 2014; Manning, 1992; Reeder, 1999). These innovations have changed the nature of policing. However, some innovations produced unintended consequences, and police failed to make effective use of others. For example, prior to the introduction of motorized patrol, police foot patrol primarily aimed to prevent crime (Wilson, 1968). The advent of motorized patrol throughout the mid-twentieth century aimed to enable more rapid response (Moskos, 2007). However, in combination with the advent of two-way radios it changed the nature of patrol. Sherman (1983) observed: “Instead of watching to prevent crime, motorized police patrol [is] a process of merely waiting to prevent crime” (p.151). Studies show an unintended consequence of these changes: much general-duties/patrol officers’ time is spent unproductively (Famega, 2005; Famega, Frank, & Mazerolle, 2005; Robinson & Hutton, 1989).

Similarly, in the latter half of the twentieth century, some police agencies introduced a single emergency phone number (Kelling & Moore, 1988; Mazerolle et al., 2002). This innovation made it easier for citizens to contact police, which increased the volume of calls that police are required to handle (Skogan & Hartnett, 1997; Sparrow et al., 1990; Walker, 1992). This increase in volume hampered the ability of police to focus on
problem solving activities. As a result some scholars argue that police should try to wean the public from telephoning police to report matters other than emergencies (Goldstein, 1990; Hawkins, 1996).

To cope with and better manage both the increasing volume of calls and the motorized patrol resources, most police agencies operate centralised call centres with computer-aided dispatch systems (Manning, 1992; Reiss, 1992; Sampson, 2004). Reiss (1992) describes this innovation as reinforcing the bureaucratic centralisation of command and control. This centralisation had the impact that “The separation of the working police from the communities policed was organisationally complete” (p.52). More recently the advent of single non-emergency phone numbers sought to alleviate demand on the emergency phone number system. A study of the impact of introducing a single non-emergency number in Baltimore, Maryland, concluded that such systems “have the potential to help police efforts to better manage citizen calls and more aptly determine appropriate police responses” (Mazerolle et al., 2002, p.119). However, this same study also noted that police failed to alter dispatch practices in a way that would capture all of the potential benefits. This finding suggests that technological innovation alone, without addressing associated process changes, may limit police effectiveness. This theme is strongly evident the literature (Colvin & Goh, 2005; Garicano & Heaton, 2007). Police often introduce technological innovations but fail to implement complementary changes to management and operational practices. As a result, increased administration and organisational complexity can be created without necessarily enhancing police effectiveness (Chan, 2001).
Collier, Edwards, and Shaw (2004) argue that cultural barriers within the policing context create resistance to change, which limits the extent to which police embrace improved ways of operating. Ashby, Irving, and Longley (2007) provide an example of this: Diffusion and adoption of *Geographic Information Systems* (GIS) by police has been slow. Weisburd and Eck (2004) argue that, despite strong evidence that geographically specific policing strategies are effective, police fail to make use of the appropriate technology. More recently, in a multi-agency study Koper, Lum, and Willis (2014) found that “police often fail to make strategically optimal uses of technology for reducing crime and serving citizens” (p. 1).

The *theory of reasoned action* posits that a person’s behaviour is influenced by their attitudes about the behaviour subjective norms (Fishbein & Ajzen, 1975; Manstead & Parker, 1995; Robey, 1979; Schultz & Slevin, 1975). This perspective provides a possible explanation of what factors influence the adoption of technology. Based on this perspective Davis (1989) developed a *technology acceptance model* which found that usefulness and perceived ease of use influence technology acceptance. However, attitudes and perceptions about technology differ among different professionals (Cork, Detmer, & Friedman, 1998). Colvin and Goh (2005) applied the technology acceptance model developed by Davis (1989) to police, and found that police attitudes are more strongly influenced by information quality and timeliness than they are by usefulness and ease of use. However, evidence is lacking on why this is so or whether this influences police effectiveness. In this thesis I seek to address this gap in the literature by examining the relationship between police attitudes, use of information, and behaviour.
The concept of police effectiveness is a central issue for my thesis. As discussed in Chapter 2, two strong theoretical themes emerge in the literature around police effectiveness. One relates to the technical-rational aims of achieving the purposes for which police is established; the other relates to achieving organisational legitimacy to secure the mandate and resources to operate. The literature on police effectiveness identifies numerous dimensions of police effectiveness within these themes. In the next section I will discuss this literature with a focus on effectiveness of police in responding to calls for service.

### 3.3 Dimensions of Police Effectiveness

Scholars such as Dobson (1982) and Maguire (2003) theorise about what constitutes police effectiveness. In doing so they concluded that effectiveness is a complex concept with many dimensions. Dobson (1982) describes one definition of effectiveness as being the degree to which an organisation achieves the purposes for which it is established. Dobson (1982) and Maguire (2003) identify other dimensions of performance, including efficiency (concerned with avoiding waste), service quality (competence and citizen/customer satisfaction), integrity (rule-breaking), accountability, and legitimacy (the extent to which citizens accept the police as proper and justified agents of the government).

Police can be effective but not have integrity by obtaining a conviction through fabricating evidence. They can be efficient but not deliver service quality by managing with limited resources and having less down-time, yet being inwardly focussed and uncommunicative with citizens. They can have legitimacy without accountability by
avoiding public scandals and by reporting measures that make them appear effective but which bear little resemblance to operational reality.

Public police are funded by government to deliver service to citizens. In this section I explore police effectiveness from the perspectives of government and citizens. Understanding these perspectives will assist in the assessment of whether police are effective in responding to calls for service.

**Symbolic justice and authoritative intervention**

A democratic state government may wish to reinforce its legitimacy as a democracy by demonstrating responsiveness to citizens (Herbert, 2006). Similarly, it will also wish to demonstrate *symbolic justice*, showing offenders and the public that a regime of law exists (Silberman, 1978). Both citizens and governments require that police demonstrate *authoritative intervention* to restore order when required (Bayley, 1994b).

**Harm prevention**

Citizens may value the feeling of personal safety engendered by confidence that someone who can prevent harm can be called on in an emergency or in a dangerous situation. If police fail to prevent harm when called, citizens can lose confidence in police. For example, in 2004 New Zealand Police was heavily criticised in the media after police failed to respond to calls for service in a way that could have prevented harm to citizens. This caused the Commissioner of Police to initiate an external review of the New Zealand Police emergency response system (Corboy et al., 2005). Police strive to avoid such crises of confidence. They seek to maintain and enhance their own organisational legitimacy, both with the public and government, thereby securing status and resources (Herbert, 2006; Lawton et al., 2000; Tyler & Huo, 2002).
Public support

In addition to these direct objectives of symbolic justice, authoritative intervention, and harm prevention, police desire to leverage public support for their efforts to combat crime. Sparrow, Moore, and Kennedy (1990) view citizens as society’s first line of defence against crime. Citizens vastly outnumber police and will often be present at or participants in situations about which police require information. The police can only reasonably expect help from the public if they give the public the help it wants. Hence delivering an effective service of responding to calls for help from the public is necessary to leverage the support police require from the public. The benefits of community policing are premised on an underlying assumption that satisfied customers of police service will become the eyes and ears of police (G. Cordner, 1995; Goldstein, 1987a; M. H. Moore, 1992; Skogan, 1994b; Wycoff, 1995).

Satisfaction with police

Studies of subjective measures of police service delivery show little direct relationship with objective measures (Brown & Coulter, 1983; Stipak, 1979). For example, Brown and Coulter (1983) conducted a telephone survey of 538 citizens in Tuscaloosa, Alabama. The survey asked respondents about their satisfaction with police services, as well as the protection, response times, and treatment of people in their neighbourhoods. Data collected were integrated with police administrative data about services provided. Brown and Coulter (1983) found that citizens’ satisfaction with police response time was independent of actual response times. In a separate study Stipak (1979) integrated responses from the Los Angeles Metropolitan Area Survey (LAMAS) on urban services (n=1,028) with administrative data. This administrative data included the services provided by police and other agencies as well as demographic characteristics of
neighbourhoods. Like the Tuscaloosa study Stipak (1979) found little evidence that service characteristics affect citizens’ evaluations of services.

A range of possible reasons exist for the lack of relationship between objective and subjective measures. Citizens may be unaware of the true level of service they receive (Brown & Coulter, 1983; Stipak, 1979). Alternatively, aggregation and statistical manipulations of objective measures may obscure or confound correlations, or it may be that subjective and objective measures may not measure the same thing (Parks, 1984). Furthermore, citizens’ beliefs about objective facts, such as the level of crime, may be influenced by their subjective perceptions of police. This was demonstrated in a telephone survey of citizens in Chicago (Skogan & Hartnett, 1997) which found that citizens’ who perceive the greatest reductions in serious crime also perceive that police were most responsive to citizens’ concerns. In a more recent Chicago telephone survey of 3,005 respondents Skogan (2006) found an asymmetrical relationship between public attitudes to police and the nature of encounters. Negative encounters engender unfavourable attitudes, but positive encounters have negligible impact. Myhill and Bradford (2011) produced a similar finding when analysing panel data from the UK National Reassurance Policing Programme.

**Attitudes to police**

Citizens’ attitudes towards police are influenced by a mix of direct personal experiences of specific encounters and generic perceptions about police in general (symbolic attitudes). For example, in a survey of 613 respondents in Cincinnati, Ohio, over half of the respondents cited specific encounters with police officers as influencing their attitude to police, yet almost one third of the respondents expressed reasons for their attitude
towards police that are unrelated to encounters with police (Frank, Smith, & Novak, 2005). A telephone survey of 509 adults in the city of Providence, Rhode Island, found that personal experience is more influential than symbolic attitudes on citizens’ perceptions of police courtesy and fairness (Orr & West, 2007). However, both personal experience and symbolic attitudes are important in influencing citizens’ perceptions of crime seriousness and overall police performance.

Numerous studies have shown that minorities appear to evaluate law enforcement authorities less favourably than the average (Albrecht & Green, 1977; Bordua & Tift, 1971; D. L. Carter, 1985; Decker & Wagner, 1981; Frank et al., 2005; Jefferson & Walker, 1993; Weitzer & Tuch, 1999). However, studies that control for confounding factors such as neighbourhood social integration, class, or media influence, have explained much of this difference (Cao et al., 1996; Dowler, 2003; Parks, 1984; Priest & Carter, 1999). For example, when analysing the 539 responses to a survey that was sent to 1,000 randomly selected residents in Cincinnati, Cao, Frank, and Cullen (1996) concluded that, although initial bivariate analysis “revealed a significant race effect, the introduction of contextual factors eliminated the relationship” (p.12).

3.4 Conclusion

In this chapter I reviewed prior research that shows police face many and complex challenges to be effective in responding to calls for service. To manage these challenges, police introduced operational and technological innovations. However, these innovations were often not fully embraced or had unintended consequences.
Police also introduced performance measurement frameworks which can help manage complexity by enabling members of an organisation to make sense of what is happening in complex organisational contexts (Gill, 2011). However, many scholars observe that police closely monitor measures related to crime but ignore the many other police activities and tasks (Fleming, 2009; Kelling, 1992; Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; M. H. Moore & Braga, 2003; Skogan & Frydl, 2004). These scholars argue that the absence of broader performance measures prevents police from being optimally effective. For example, Brunetto and Farr-Wharton (2005) and Lawton et al. (2000) argue that the absence of broader measures prevents police executives from motivating their middle managers and front-line staff to change their approach to policing to be more responsive to the concerns of the communities that they serve. Maguire (2003) and Sparrow et al. (1990) describe how police developed a strong identity as society’s professional crime fighters as a result of the professionalization reforms of the twentieth century. They argue that providing general assistance to citizens does not fit with this identity and, as a result, police are less responsive to miscellaneous calls for service. This narrow focus, they suggest, lessens the relationship the police have with the community which, in turn, reduces the willingness of citizens to actively support the police. However, creating a broad and effective performance measurement framework is challenging (Telep, 2014). Legrand and Bronitt (2012) suggest numerous broader measures of police effectiveness, but argue that “Measuring these is not a trivial exercise and raises difficult methodological questions” (p. 15).

In this thesis I build on the literature reviewed in Chapters 2 and 3 by exploring whether performance measurement frameworks in Queensland Police Service and New Zealand Police have a narrow focus on crime and, if so, whether this limits the effectiveness of
Specifically, scholars such as Maguire (2003), Reiss (1992), and Sparrow et al. (1990) suggest that police are less effective than they could be because police have developed a narrow identity as society’s crime fighters which is reflected in a commensurate narrow focus in the performance measures monitored by police. They explain this as being a consequence of how policing evolved during the twentieth century. However, they provide limited empirical evidence to prove that police are ineffective. In this thesis I seek to develop evidence to test these arguments and extend our theoretical understanding of police effectiveness in responding to calls for service. In Study 1 I draw on the literature reviewed in Chapters 2 and 3 to frame questions about performance measurement in relation to delivering a service of responding to calls for service from citizens.
Chapter 4: Study 1 Method

4.1 Research Questions

This thesis seeks to determine whether police are effective in responding to calls for service from citizens. Technical-rational theory considers effectiveness in terms of how well police deliver the service of responding to calls for service (Willis et al., 2007). Institutional theory considers effectiveness in terms of how well police reflect the values of their constituents (Meyer & Rowan, 1977). Using these perspectives, in Study 1 I explore the attitudes and behaviours of police personnel involved in delivering a service of responding to calls for service from citizens.

As part of the public sector management reforms of the 1980s and 90s, technical-rational theory guided the introduction of performance measurement frameworks into police agencies (Hoque et al., 2004; O'Malley & Hutchinson, 2007; Palmer, 1997; Ritsert & Pekar, 2009; Willis et al., 2007). These reforms sought to improve efficiency and effectiveness of service delivery. However, review of previous research reveals that significant gaps in knowledge exist regarding the use and impact of performance measurement in the policing context and those factors that influence police effectiveness in responding to calls for service from citizens. Specifically, three questions arise from the literature reviewed in the previous chapters.

Research Question 1. How do police managers use performance information, and do they use performance information effectively?
Scholars such as Maguire (2003), Reiss (1992), and Sparrow et al. (1990) suggest that police are less effective than they could be because they focus too narrowly on performance measures related to crime reduction. Coleman (2008) suggests that police managers may lack the skills to use performance information effectively, and Andersson and Tengblad (2009) suggest that police managers are so socialised into thinking like police officers that they are unable to think like managers. However, evidence supporting these ideas is limited.

**Research Question 2. Does performance measurement help or hinder police effectiveness in either technical-rational or institutional terms?**

Police officers exhibit a high level of commitment to their job (Brunetto & Farr-Wharton, 2003), so may use performance measurement to improve service delivery by improving their management of processes and resources. However, police agencies are strongly institutionalised (Crank, 2003). So, it may be that police implement performance measurement frameworks to gain organisational legitimacy by appearing progressive, rather than to improve service delivery (Taylor, 2009; Willis et al., 2007). In either case, prior research presents mixed findings on the use and impact of performance measurement in police (Ashby et al., 2007; Bayley, 2008; Hoque et al., 2004; M. H. Moore & Braga, 2003; Savage, 2007).

**Research Question 3. What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?**
Evidence exists of loose coupling within police (Collier, 2001; Lawton et al., 2000; Meyer & Rowan, 1977). However, prior research has not adequately explored the nature of loose coupling in police organisations and its influence on police effectiveness, nor whether performance measurement influences loose coupling in police organisations.

I seek to answer these three research questions in the context of police response to calls for service in New Zealand and Queensland. To address the research questions I use a mixed methods research design that includes two interconnected studies. Considering the current state of knowledge, my first study is exploratory. An exploratory approach allows me to build upon and clarify the propositions put forward in prior research, as well as develop new ideas about the factors impacting police effectiveness in responding to calls for service. This approach also allows me to consider the complex interplay between police use of performance information, loose coupling, and the effectiveness of police in both technical-rational and institutional terms. Furthermore, an exploratory approach assists in generating hypotheses about the impact of performance measurement on the effectiveness of police in responding to calls for service. I use the findings of Study 1 to generate hypotheses about the relationship between performance measurement and police effectiveness in both technical-rational and institutional terms. These hypotheses are tested by undertaking further quantitative analysis in Study 2 (see chapters 6 and 7).

In Study 1, to investigate the relationship between police perspectives, behaviours, and effectiveness, I conduct semi-structured in-depth interviews with police personnel in New Zealand and Queensland who are involved with delivering a service of responding to calls for service from citizens. My goal from these interviews is to generate a better understanding of the relationship between the perspectives of police personnel and the
effectiveness of police. In designing these interviews I draw on three theoretical perspectives that emerge from the literature: technical-rational theory, institutional theory, and the theory of loose coupling. In the interviews I also explore the associated issue of police use of performance information – how and why police managers use performance information, and what influence this has on the effectiveness of police in responding to calls for service from citizens.

This chapter describes the research methods used in Study 1. In particular, I explain how and why the research sites and interviewees are selected. I then describe the theoretical approach taken in designing the interviews, the interview procedure, and the structure of the interviews. Finally, I describe how data are analysed.

4.2 Interview Sites

Most prior research reviewed in the previous two chapters concerns studies conducted in Europe and North America. However, the questions arising from prior research are not only relevant to those continents, but also to Australia and New Zealand which were also strongly influenced by NPM reforms (Gill, 2011; Jesson, 1989). To develop an understanding of the relationships between performance measurement, loose coupling, and police effectiveness in responding to calls for service, I conduct studies in Australia and New Zealand.

Australia and New Zealand share a similar culture and have similar police functions. New Zealand has a single national police agency. Australia has a number of states and territories, each with their own state government and police agency. Table 4.1 presents
figures as at 30 June 2009 for the normally residential population of each state and the number of police personnel in each state.

Table 4.1
State and Police Populations

<table>
<thead>
<tr>
<th>State</th>
<th>Population (million)</th>
<th>Police staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Zealand</td>
<td>4.32(^2)</td>
<td>11,881(^3)</td>
</tr>
<tr>
<td>Queensland</td>
<td>4.35(^4)</td>
<td>14,222</td>
</tr>
<tr>
<td>New South Wales</td>
<td>7.04</td>
<td>19,153</td>
</tr>
<tr>
<td>Victoria</td>
<td>5.37</td>
<td>13,901</td>
</tr>
<tr>
<td>Western Australia</td>
<td>2.20</td>
<td>7,474</td>
</tr>
<tr>
<td>South Australia</td>
<td>1.61</td>
<td>5,431</td>
</tr>
<tr>
<td>Tasmania</td>
<td>0.50</td>
<td>1,602</td>
</tr>
<tr>
<td>Northern Territories</td>
<td>0.22</td>
<td>1,587</td>
</tr>
<tr>
<td>Australian Capital Territory</td>
<td>0.35</td>
<td>945</td>
</tr>
</tbody>
</table>

As shown in Table 4.1, Queensland has a similarly sized population and similarly sized state police agency to New Zealand. New Zealand Police and Queensland Police Service also provide a similar response service to calls for service from citizens. These similarities combine to minimise the potential for confounding environmental factors. Using two similar police organisations also enabled me to compare within and across organisations. This provided opportunities to examine the generalisability of the findings.

\(^2\) Statistics New Zealand

\(^3\) New Zealand Police Annual Report 2008-09 p. 79

\(^4\) All figures for Australian state and territories sourced from the Report on Government Services 2010
At the time that this study was undertaken\(^5\), both New Zealand Police and Queensland Police Service were organised into police districts. In New Zealand there were 12 police districts, and in Queensland there were 31 police districts. In charge of each district was a *district commander* (New Zealand Police) or *district officer* (Queensland Police Service), most of whom held the rank of Superintendent. Each district was further subdivided into a number of *areas* (New Zealand Police) or *divisions* (Queensland Police Service), each with a commander. Most commanders of areas/divisions held the rank of inspector.

Both New Zealand Police and Queensland Police Service operate *computer aided dispatch* (CAD) systems in which information collected during jobs of responding to calls for service is recorded. New Zealand Police operates a single CAD system, known as CARD, and has three *police communications centres* (PCCs) across New Zealand. These centres receive calls and dispatch police units to attend jobs. Each PCC is responsible for dispatching jobs to a different group of districts. I selected a district from each of these groups, thereby ensuring all three PCCs were included.

At the time I conducted Study 1, some Queensland Police Service districts used a CAD system that was similar, but not identical, to the system used by New Zealand Police. Other Queensland Police Service districts used either an older, more basic IT system, or managed their dispatch manually. For comparability with New Zealand Police, and to ensure availability of data for Study 2 (see Chapter 6), I selected three districts from Queensland that used the CAD system and a different PCC from each other. These districts

---

\(^5\) Since I undertook Study 1, Queensland Police Service has implemented structural changes, altering some district boundaries and staffing, and creating a centralised call centre known as Policelink, to handle non-emergency calls for service. Also, both organisations continually update the technologies they use. In this chapter I describe the technologies and organisations as they existed at the time this study was undertaken.
districts were all predominantly metropolitan. So, to maintain comparability, I also selected three predominantly metropolitan districts in New Zealand. To protect confidentiality of interviewees, the specific districts included in this study are not identified.

4.3 Participants

4.3.1 Ranks and roles of participants
Interviewees were selected from each police headquarters and the six districts using a purposive sample. This ensured that all ranks and roles involved in responding to calls for service were included in this research. Six roles were identified: Communications room operator, general duties officer, PCC manager, district manager, senior executive, and performance manager.

Communications Room Operators
Communications room operators work in PCCs answering phone calls from citizens and dispatching police units to attend jobs. Most communications room operators are civilian police employees. However, a small proportion are police officers.

General Duties Officers
This group consists of general duties constables and senior constables working in districts. A core part of their role involves responding to calls for service by attending the scene when dispatched by a communications room operator. A unit consisting of one or two general duties officers in a car is the most common type of unit dispatched by communications room operators.
**PCC Managers**

This group consists of supervisory or managerial police staff working within or in command of PCCs. The most junior personnel in this group are first line supervisors. In Queensland Police Service these first line supervisors are all police officers, usually having a rank of sergeant. However, in New Zealand Police first line supervisors may be sergeants but can also be civilian personnel. These civilian supervisors hold the same seniority and pay band as their police officer counterparts.

In both New Zealand Police and Queensland Police Service more senior management positions in this group are all held by police officers ranging in rank from senior sergeant to superintendent, depending on their role in the PCC and the size of the PCC. One difference between New Zealand Police and Queensland Police Service is that New Zealand Police PCCs do not have senior sergeants in middle management roles. PCC staff in both agencies are rostered onto shifts. However, New Zealand Police only has three large PCCs, whereas Queensland Police Service has a larger number of PCCs which range in size. Consequently, managers in charge of shifts in New Zealand Police are all inspectors. However, PCCs in Queensland employ a number of senior sergeants.

**District Managers**

This group consists of supervisory or managerial police staff with direct line control of general duties officers or their supervisors or managers. All participants in this group are police officers based in districts. They include first line supervisors (all sergeants), officers in charge of police stations (all senior sergeants), officers in charge of areas (New Zealand Police) or divisions (Queensland Police Service) (all inspectors), and officers in command of districts (all superintendents). The manager in command of a PCC, district
or division/area was automatically selected by virtue of the PCC, district, or division/area they command being included in the study. Senior executives and performance managers were fewer in number, so required application of a different method of selection.

### Senior Executives

This group represents the most senior executives in main headquarters of the agency. It contains the smallest number of participants (n=8), as both New Zealand Police and Queensland Police Service have fewer senior executives than personnel in other roles. Although relatively small, this group is still representative as it contains almost all personnel in the agency that fit the criteria of (a) having a rank of at least assistant commissioner, (b) being based at the main headquarters of the agency, and (c) having responsibility for some aspect of operating or managing police response to calls for service. It contains one of the two commissioners, three of the four deputy commissioners, and four assistant commissioners in roles with responsibility for some aspect of managing police response to calls for service. To enlarge this group would either compromise the anonymity of participants or introduce positions of lower seniority or less relevance to the desired role.

### Performance Managers

Participants in the above five groups are directly involved in delivering or managing police response to calls for service. I also interviewed nine police personnel with more general responsibilities for performance measurement. Interviewing performance managers provides an insight into the performance measurement frameworks used by the agencies. This group contains a mix of senior police officers and senior civilian personnel. Although relatively small, this group is still representative as it contains almost all
personnel with responsibility for the agency’s OPR (see Section 5.3.3) or formal performance reporting framework.

In selecting participants to interview in Study 1 I sought to select a representative cross-section of police personnel across three dimensions: (a) organisation (New Zealand Police and Queensland Police Service), (b) role (six types of role), and (c) seniority. To adequately represent these three dimensions I interviewed 121 personnel from two police agencies: New Zealand Police (n=63) and Queensland Police Service (n=58). The personnel interviewed were employed in roles associated with delivering or managing a service of responding to calls for service from citizens.

Table 4.2 presents the numbers of police personnel interviewed from each rank or equivalent seniority level. To preserve anonymity of participants, in this table the ranks of commissioner and deputy commissioner are merged. I was fortunate to be able to interview the commissioner and one of the two deputy commissioners in one agency, and both deputy commissioners in the other agency. Participants are also grouped into six types of role. Table 4.3 presents the distribution of participants across these six roles and two agencies.
Table 4.2
Interviewees by Rank

<table>
<thead>
<tr>
<th>Rank equivalent</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Officer</td>
<td>Civilian</td>
<td>Officer</td>
</tr>
<tr>
<td>More junior than constable</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Constable/ Senior Constable</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Sergeant</td>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Senior Sergeant</td>
<td>6</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Inspector</td>
<td>9</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Superintendent</td>
<td>7</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Assistant Commissioner</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Commissioner/ Deputy Commissioner</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>46</td>
<td>17</td>
<td>45</td>
</tr>
</tbody>
</table>

Table 4.3
Interviewees by Role

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications room operator</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>General duties officer</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>PCC manager</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>District manager</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Senior executive</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Performance manager</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>63</td>
<td>58</td>
</tr>
</tbody>
</table>

The groupings by rank in Table 4.2 align closely with the different types of role in Table 4.3. Interviewees more junior than constable are all civilian communications room.
operators; constables and senior constables are all non-management police officers, most of whom are general duties officers; all sergeants are first-line supervisors, either in districts or PCCs; most senior sergeants are either managers in charge of police stations or managers in PCCs; most inspectors are either managers in command of police areas/divisions or managers in PCCs; the superintendents are spread across roles in headquarters or in command of a district or one or more PCCs; the commissioner, deputy commissioners and assistant commissioners are all senior executives at the agency’s main headquarters. In presenting results, all interviewees in PCCs with a rank of sergeant to superintendent are treated as a single group called PCC managers; all interviewees in districts with a rank of sergeant to superintendent are treated as a single group called district managers. This approach ensured at least ten interviewees were selected from each of the four largest groups in each agency, and that interviewees were selected from each rank and role in each PCC and district. This produced slightly more interviewees who are district managers, due to the districts in both agencies having four distinct layers of management.

4.3.2 Procedure used to select participants

Each police agency nominated a liaison officer for the project. This liaison officer assisted in arranging access to interviewees. Due to the complexity of staff rosters and lack of a suitable centralised database, apart from personnel in the two smallest groups (senior executives and performance managers), it was not feasible to apply a randomised sampling algorithm to select people to interview from a list of all relevant police personnel. Instead, participant selection occurred based on role responsibilities. The following procedure was used to select and interview participants:
Through the liaison officers appointed by each agency, I obtained the permission of senior managers to arrive in person at various police work locations and select staff to interview. At each police work location I identified to the workplace supervisor when I was ready to conduct the next interview. The next person to come free from a job, if not urgently required to attend another job, was then asked if they would speak to me in an interview room on the same floor of the building they were working in. Although the selected participants represent a sample of convenience rather than a mathematically randomised sample, the selection process ensured that I interviewed a cross-section of staff in each role rather than staff selected by the organisation. By taking the structured selection approach that I did, I was able—at least in part—to mitigate against selection bias due to subjectivity of the researcher or managers in the police agency being studied.

**Selection of participants based in police headquarters**

Potential interviewees from the two smallest groups: Senior Executives (n=8) and Performance Managers (n=9) all worked in the state headquarters and were few in number. Therefore they were easy to identify from discussions with the liaison officers. To protect the anonymity of interviewees, for each of these groups within each agency all potential participants except one were interviewed. This resulted in four senior executives from each agency being interviewed, but six Performance Managers from New Zealand Police and only three from Queensland Police Service being interviewed. This difference reflects slightly different organisational structures and recent personnel changes in each agency.
Selection of participants not based in police headquarters

As discussed earlier in this chapter, I selected personnel from three districts within each of Queensland Police Service and New Zealand Police, and from the PCCs servicing these districts. All of New Zealand Police uses the same computer-aided dispatch system (CARD). However, the equivalent type of system in Queensland Police Service (CAD) is only used in some areas. I selected districts from Queensland Police Service that use the CAD system. This ensured the operating environments of participants from each agency are as similar as possible, and ensured comparable data would be available for Study 2.

Procedure to gain agreement of participants

I conducted all interviews in a private interview room or office. In the initial stages of the interview I provided information to the potential participant and sought their consent to be interviewed, as required by Griffith University ethics policies\(^6\). I gave the potential interviewee the options to decline to be interviewed, to be interviewed outside the workplace where their employer would not be aware of this, or to be interviewed at that time. To ensure no coercion by the police agencies the person’s supervisor and more senior managers were not told whether the individual agreed to be interviewed. The full informed consent package is provided in Appendix A.

Only one person out of 122 asked declined to be interviewed. This person declined the interview on the basis that they were too busy to be interviewed. I conducted all interviews between mid-October 2008 and mid-February 2009.

\(^6\) Ethics approval reference number CCJ/03/08/HREC
4.4 Semi-structured Interviews

I conducted semi-structured interviews to address the research questions:

RQ 1. How do police managers use performance information, and do they use performance information effectively?

RQ 2. Does performance measurement help or hinder police effectiveness in either technical-rational or institutional terms?

RQ 3. What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?

In developing the interview schedules, I applied Pawson’s (1996) theory-driven approach in which “the researcher’s theory is the subject matter of the interview, and the subject is there to confirm or falsify and, above all, to refine that theory” (p. 299). In this way, I approached the interviews with the assumption that there is a relationship between police perspectives, behaviours and effectiveness, and that contextual features of the police organisation environment influence this relationship. I sought to confirm this relationship and refine my understanding about this relationship. In line with this approach, interview questions were constructed around the themes identified in the literature. For example, perceptions and evidence of problems in measuring performance were explored (Brunetto & Farr-Wharton, 2003, 2005; Lawton et al., 2000).

The interviews involved my asking open-ended questions. This approach allowed the participant to elaborate on a response, and allowed me to probe for greater depth or to
seek clarification of a response to a question. The interviews also explored the reasons underpinning the specific views reported by the interviewee.

**Interview schedules**

To ensure the relevance of questions to the different roles of participants, four different schedules were developed. Each schedule contains 26 primary questions. Additionally, the schedules contain a number of secondary questions prompting the participant to clarify or expand on their answer to the relevant primary question. The full schedules are provided in Appendix B and summarised in Table 4.4. These four schedules are as follows:

- *Frontline interview schedule* – Used to interview communications room operators and general duties officers
- *Middle managers interview schedule* – Used to interview PCC managers and district managers
- *Senior executives interview schedule* – Used to interview senior executives
- *Performance managers interview schedule* – Used to interview performance managers

All schedules commence with an identical preamble by the interviewer to ensure all participants were given a common understanding of the context about which the interview questions relate. This preamble reads as follows:

In this interview I’d like to concentrate in particular on police response to calls for service from the public. This means all the things that police do from when a member of the public either requests help from police or reports to police a matter
requiring police action, through to the end of police initial attendance at the scene should police decide this is required. This excludes subsequent investigations, prosecution of alleged offenders, crime prevention initiatives, or other functions not related to police initial response to calls for service from the public. Do you understand what I mean, or would you like me to clarify any of this?

Table 4.4
Summary of Interview Schedules

<table>
<thead>
<tr>
<th>Question</th>
<th>Schedule</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F,M,S,</td>
<td>Identify jobs about which questions will be asked.</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>Screener question, to confirm role of interviewee.</td>
</tr>
<tr>
<td>3</td>
<td>F,M,S</td>
<td>Identify what interviewee perceives as success/ effectiveness at a job level.</td>
</tr>
<tr>
<td>4</td>
<td>F,M,S,P</td>
<td>Discover interviewee’s awareness of performance measurement.</td>
</tr>
<tr>
<td>5</td>
<td>F,M,S,P</td>
<td>Identify why police do not measure performance (if relevant).</td>
</tr>
<tr>
<td>6</td>
<td>F,M,S,P</td>
<td>Identify purpose of performance measurement.</td>
</tr>
<tr>
<td>7</td>
<td>P</td>
<td>Identify how police measure performance.</td>
</tr>
<tr>
<td>8</td>
<td>F,M,S,P</td>
<td>Identify performance measures.</td>
</tr>
<tr>
<td>9</td>
<td>F,M,S,P</td>
<td>Awareness of who uses performance reporting.</td>
</tr>
<tr>
<td>10</td>
<td>F,M,S,P</td>
<td>Awareness of purposes of performance reporting.</td>
</tr>
<tr>
<td>11</td>
<td>F,M,S,P</td>
<td>Perspective on what is important for police effectiveness.</td>
</tr>
<tr>
<td>12</td>
<td>F,M,S,P</td>
<td>Identify how interviewee monitors effectiveness.</td>
</tr>
<tr>
<td>13</td>
<td>M,S,P</td>
<td>Identify interviewee’s preference for monitoring or measuring effectiveness.</td>
</tr>
<tr>
<td>14</td>
<td>M,S,P</td>
<td>Identify interviewee’s frustrations and limitations with measurement.</td>
</tr>
</tbody>
</table>

7 In Table 4.4 the purpose of each question is given; the schedules containing each question are also indicated. Schedules are coded in Table 4.4 as follows: F = frontline, M = middle managers, S = senior executives, P = performance managers.
<table>
<thead>
<tr>
<th>Question</th>
<th>Schedule</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>F,M,S,P</td>
<td>Identify a time when police performance was poor.</td>
</tr>
<tr>
<td>16</td>
<td>F,M,S,P</td>
<td>Identify consequence of poor performance.</td>
</tr>
<tr>
<td>17</td>
<td>F,M</td>
<td>Discover interviewee’s personal involvement in performance reporting.</td>
</tr>
<tr>
<td>18</td>
<td>F,M</td>
<td>Discover interviewee’s awareness that what they record is used in performance measurement.</td>
</tr>
<tr>
<td>19</td>
<td>F,M,S,P</td>
<td>Does interviewee believe performance monitoring is helpful, and how has it affected the interviewee?</td>
</tr>
<tr>
<td>20</td>
<td>P</td>
<td>Identify problems encountered in monitoring performance.</td>
</tr>
<tr>
<td>21</td>
<td>P</td>
<td>Identify changes made to performance frameworks because of problems in measuring performance.</td>
</tr>
<tr>
<td>22</td>
<td>F,M,S</td>
<td>Does interviewee believe his/her perceptions about performance monitoring influence what he/she records?</td>
</tr>
<tr>
<td>23</td>
<td>M,S</td>
<td>Discover what, if anything, managers monitor in addition to the formal performance measurement frameworks.</td>
</tr>
<tr>
<td>24</td>
<td>F,M,S</td>
<td>Interviewee’s perspective on how his/her performance is assessed.</td>
</tr>
<tr>
<td>25</td>
<td>F,M,S</td>
<td>Identify what the interviewee does to please his/her supervisor.</td>
</tr>
<tr>
<td>26</td>
<td>F,M,S,P</td>
<td>Open question, giving opportunity for unprompted comment.</td>
</tr>
</tbody>
</table>

Where relevant an identical question is repeated in different schedules. Not all questions are relevant to all groups of participants. Differences between the four interview schedules are summarised in Table 4.4. Overall, the questions were designed to serve eight broad purposes. These are as follows:

- Screen participants to confirm that their role aligns with one of the groups being studied, and that the appropriate schedule is being used
- Identify the factors that determine whether the participant perceives a given instance (job) of responding to a call for service as being successful
• Identify more generally what the participant believes constitutes effectiveness and good performance in police responding to calls for service
• Discover the participant’s level of awareness and knowledge of the purposes and mechanisms used within their organisation to measure performance
• Explore the participant’s attitudes to and use of performance information, and how performance information influences their actions
• Identify what occurs in the participant’s organisation when performance is identified as unsatisfactory
• Identify what factors influence decisions about which performance information to report
• Identify what participants do to ensure their superiors are satisfied with their performance

**Differences between the four schedules**

The frontline interview schedule frames questions about specific instances of responding to calls for service in which the participant had personally been involved. This schedule omits questions about management tasks that the interviewee would not have personally performed as a communications room operator or general duties officer.

The middle managers schedule includes questions about management actions relevant to the research questions. It also asks similar questions about response jobs to those in the frontline schedule. However, questions are modified slightly to refer to jobs involving the participant’s staff, rather than jobs the participant did personally. This modification reflects the difference in role between the frontline groups and the middle manager groups. Making these changes ensures the questions remain relevant to each role. For
example, Question 1 in the frontline schedule reads: “What were some of the jobs you
did during the last few weeks that are typical of the sort of work you do in your current
position?” Question 1 in the middle manager schedule reads: “Do you have knowledge
about any specific jobs involving police response to calls for service from the public that
have occurred in the last few weeks, and that involved either staff under your command
or the area that you are responsible for?” When middle managers being interviewed
indicated that they did have knowledge of such jobs, I prompted them to identify specific
elements.

The senior executives schedule includes most of the questions asked in the middle
managers schedule, but reframes these to refer more generally to the entire organisation
rather than the particular work group that a middle manager would have responsibility
for. For example, as given above, Question 1 in the middle managers schedule asks about
the participant’s knowledge of jobs “that involved either staff under your command or in
the area that you are responsible for.” However, the senior executives schedule asks about
the participant’s knowledge of jobs that had occurred within the last few weeks within
the participant’s police agency as a whole. Similarly, Question 24 in the middle managers
schedule asks: “How do you know if your superiors think your performance is good?”
Whereas question 24 in the senior executives schedule asks: “How do you know if the
minister of police and other key stakeholders think your performance is good?”

The performance managers schedule omits questions about specific jobs of responding to
calls for service, and about the participant’s behaviour in reporting the performance of
their work group. However, it includes the more general questions about performance
measurement and reporting, as well as some additional questions about the organisation’s
performance frameworks themselves. For example, questions 20 and 21 only appear in the performance managers schedule. Question 20 reads: “What problems have you encountered in monitoring police performance in responding to calls for service from the public?” Question 21 reads: “Are you aware of changes that have been made to performance measures or to the performance measurement framework because of previous problems in measuring or reporting performance?” When performance managers indicated that they did have knowledge of such problems, I prompted them to identify what these problems were, and to describe what changes were made as a result.

4.5 Analytic Approach

I employed Pawson’s (1996) theory-driven approach to develop interview questions and conduct the analysis. This theory-driven approach influences both the design of the interviews (see Section 4.4) and the analytic approach. I followed a structured process to organise and analyse the interview data. In this process I commenced by reading about a quarter of the interview transcripts to develop a sense of the whole. I then began systematically analysing the interview transcripts to identify topics within the interviews. I did not review content at this stage, but just listed topics identified from each interview transcript. I then compared the topics identified from individual transcripts, merging similar topics to produce a rationalised list. I grouped topics in this list into broader themes associated with each of the three research questions:

RQ 1. How do police managers use performance information, and do they use performance information effectively?
RQ 2. Does performance measurement help or hinder police effectiveness in either technical-rational or institutional terms?

RQ 3. What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?

Guided by the questions in the interview schedules together with the responses given by participants I identified research sub-questions related to each topic. I then coded the topics. Finally, I reviewed interviews according to the topics, elaborating a small set of generalisations to answer the research questions and sub-questions. In elaborating these generalisations I considered how they relate to the research questions and the literature through existing and new theories. For example, although I found considerable evidence of police performance measurement activity, I found little evidence of participants’ decisions being influenced by the measures being monitored. Instead, qualitative checking of individual instances appears for compliance with a standard appears to be the dominant mechanism by which performance is measured.

Chapter 5 presents and elaborates on this and other findings. From these findings I develop hypotheses that are tested in Study 2 using a quantitative method. This mixed methods approach enables me to both develop and test theoretical propositions arising from the literature and my studies. In Chapter 8 I discuss these propositions and their implications for policy, practice, and future research.
Chapter 5. Study 1 Findings

5.1 Introduction

Prior research does not provide a clear picture of whether performance measurement helps or hinders police effectiveness. Theory behind NPM suggests performance measurement can improve police performance in delivering service to citizens (O'Flynn, 2007). However, institutional theory suggests that pressure to reflect the values of, and gain legitimacy with, constituents may influence why and how police use performance measurement (Taylor, 2009; Willis et al., 2007). Loose coupling theory suggests performance measurement may actually be counter-productive, driving staff disengagement and even subversion of policy or the performance that is reported (Brunetto & Farr-Wharton, 2003; Lawton et al., 2000). Some scholars suggest police managers may lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008). However, evidence is lacking on whether this suggestion is true and, if so, why. I therefore designed Study 1 to explore these theoretical perspectives on the relationship between performance measurement and police effectiveness. The primary aim of Study 1 was to generate testable hypotheses about this relationship through qualitative interviews in two police agencies.

In Chapter 4 I identified three research questions that arise from the literature reviewed in Chapters 2 and 3:

RQ 1. How do police managers use performance information, and do they use performance information effectively?
RQ 2. Does performance measurement help or hinder police effectiveness in either technical-rational or institutional terms?

RQ 3. What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?

In this chapter, I address these three research questions. I begin the chapter by exploring how police managers make decisions affecting police effectiveness, and why they make decisions this way. I then explore how performance measurement is practised within police agencies, and what influence performance measurement has on police effectiveness. The interviews also provide an understanding of loose coupling within police agencies: what form loose coupling takes and how this loose coupling influences police effectiveness. The key themes and ideas generated through Study 1 assist in constructing hypotheses about the relationship between performance measurement and police effectiveness. The chapter concludes by developing these hypotheses.

5.2 RQ1 How do Police Managers use Performance Information, and do They use Performance Information Effectively?

When questioning interviewees about their understandings of performance measurement in police a theme emerged on how police managers use performance information. Within this theme I identified three topics: the types of performance information police managers use, why police managers make only limited use of quantitative performance information, and how police managers use quantitative performance information to influence decisions. In this section I discuss findings for these topics in relation to arguments made
in the literature about the ability of police to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008).

5.2.1 RQ1.1 What types of performance information do police managers use?

The interview transcripts reveal two distinct ways that interviewees use information to assess police performance and effectiveness: monitoring patterns in quantitative information to recognise emerging problems or trends, and reviewing qualitative details about individual jobs to determine if the actions taken by police were appropriate. Some managers demonstrated an ability to integrate both quantitative and qualitative information. Some do this better than others. However, the managers interviewed predominantly focus on managing individual jobs, and gave little evidence of interest or ability in interpreting quantitative information. For example, when asked “What do you find most helpful to monitor or measure?” one PCC manager responded: “I make a point of reading every job during the shift” [NP101]. When prompted to identify the performance measures police use to measure their performance in responding to calls for service, one senior sergeant in charge of a police station described how supervisors and more senior officers check patrol logs, occurrence sheets, and crime reports, to check that each response was appropriate [QD021]. Another senior sergeant in charge of a police station said: “I assume there is [sic] ways they monitor our CARD system that we use for jobs, but personally it’s just what I hear on the street, and the events that I have read in the paper, and stuff about our responses, or generally it’s the lack of response that appears

8 The index given in square brackets following a quotation is a unique code assigned to the interviewee who is quoted. The first letter in this index identifies the agency (N = New Zealand Police; Q = Queensland Police Service). The second letter in this index identifies the role-type (C = communications room operator; G = general duties officer; P = PCC manager; D = district manager; S = senior executive; M = performance manager). The last three digits of the index give the unique sequence number allocated to the interviewee.
in the papers” [ND090]. An inspector in a district did not believe police measure their performance in responding to calls for service because doing so would be too complex. This inspector stated: “We tend to measure by exception. So, when things go wrong we look at it that way, and if things aren’t going wrong we must be doing OK I suppose – very poor system” [ND084]. One performance manager expressed concern that police fail to make use of performance information that would enable them to avoid crises of organisational legitimacy. Instead:

One of the great indicators of performance is when things go really badly wrong. [It is] way too late by then. Priority Policing is a classic example of that. We should never have had a need for that you know, but that's one of the things police do, they wait for things to go bad before that happens. [QM007]

When discussing how performance is actually managed, interviewees readily described qualitative checking of individual jobs by an individual’s supervisor and by the supervisor’s managers. This style of performance management is evident at all levels of management from first-line supervisors to senior executives. Although quantitative measures are prevalent in performance reports at more senior levels, many managers rely on historical behaviours that were developed as a first-line supervisor. There was little evidence that these managers use quantitative performance information well to influence their decisions. Many interviewees simply assume that they are performing well unless others tell them things are not working well. For example, responses to the question “How do you know if your superiors think your performance is good?” included: “I think it’s a case of no news is good news” [NG065]; “My superior will come and tell me if my performance is substandard” [QP030]; “You get told if you’re doing a bad job, and if
you’re not given any bad feedback, well then you must be doing all right” [NC104]; “If it’s alright they won’t say anything. Again, if you’re not doing a job, you get hauled aside and told what you’re doing wrong. If you’re doing a good job you very rarely get told” [QD026].

5.2.2 RQ1.2 Why do police managers make only limited use of quantitative performance information?

In analysing transcripts I sought evidence to test ideas suggested in previous research. Coleman (2008) suggests that performance measurement in police has limited impact because police managers lack the necessary skills to use performance information effectively. Andersson and Tengblad (2009) suggest that police managers have been so socialised into thinking like police officers that they are unable to think like managers.

Managers interviewed in Study 1 gave little evidence of quantitative performance information influencing their decisions, and little evidence of skills in interpreting even basic patterns in data. For example, one inspector in a district believed that monitoring performance measures helps create a sense of sharing a common goal within police. However, this interviewee gave no evidence of actually using performance measures to influence decisions [QD056]. Some interviewees said that every call for service was so different that police actions should be considered on a case by case basis. For example, one superintendent in charge of a district stated: “We don’t produce [performance reports] for external situations, and I think with good reason, because you’ve got to look at each one of these particular cases [sic] has to be assessed on its merits” [QD017]. Similarly, a communications room operator said that calls for service are “always different, so how you can monitor and manage that performance and make sure it’s excellent without
wasting the resources is a—Maybe Socrates could help” [NC082]. One performance manager drew an analogy to measuring “how long a piece of string is. One job may take a couple of seconds on the phone call, but another job may also take hours at that initial attendance at the scene. …..It’s not like counting widgets; it’s comparing many different things” [NM062]. These examples provide one reason why police make limited use of quantitative performance information: Some police personnel believe that calls for service are so different from each other that quantifying information about them has little value.

Police managers may be reluctant to use quantitative performance information because they perceive quantitative information to be inaccurate or misused (Lawton et al., 2000, p. 17). For example, when asked what problems exist in monitoring police performance, one performance manager responded: “Data errors is a common issue” [NM120]. A PCC manager stated: “So, we’ve just got to be careful around that because with statistics, you know people can bend them any way they like to for their own purposes” [NP118]. A first-line supervisor in a district stated: “A lot of times there’s stats, stats, and more lies, and stats. So, again I think a lot of them can be—not manipulated—but can be worded to someone’s advantage” [QD026]. Similarly, another first-line supervisor in a district stated:

I have had a case where the district commander has said to me ‘Your section is not performing as well as another section; tell me why’, and he has presented me with statistics. I have presented him with my own and said ‘you are wrong’. So, there was a number of times in the past where his statistics didn’t marry up with mine. [ND069]
5.2.3 RQ1.3 How do police managers use quantitative performance information to influence decisions?

A small number of examples were found of quantitative performance information being used in a structured way. For example, New Zealand Police routinely monitors response-time performance and the time taken in PCCs to answer phone calls. Where monitoring of this nature occurs in Queensland Police Service it is typically less formal and less standardised across the whole organisation.

When asked about what performance reports are produced regarding police response to calls for service, some interviewees from Queensland Police Service identified a routine performance report of the top ten locations for repeat calls for service. This report contains both a list and counts of repeat calls for service to the same address. However, no interviewees gave any evidence that the number of repeat calls for service is monitored over time to determine if the volume of repeat calls for service is changing, or if police are becoming more or less effective in solving problems that lead to repeat calls for service.

The predominant use of quantitative information described by interviewees from both agencies is monitoring of inputs (demand volumes) to identify when workload is high. In particular, interviewees frequently said that monitoring demand volumes is helpful for allocating staffing levels. For example, one communications room operator said “it helps by establishing trends of when our busiest periods are, so we can put more staff on” [QC008]. Some interviewees, particularly in PCCs, referred to the volume of calls for service as a measure of demand. However, the quantitative information most commonly referred to is crime statistics. Both agencies measure volumes of recorded and resolved
crimes. Many interviewees perceive crime statistics as measures of demand volume. For example, a senior sergeant in charge of a police station stated: “We look at crime trends over the previous 12 months, 6 months, and 28 days, and then daily we look at crime trends and calls for service and develop strategies to address any problems” [QD019]. Some interviewees perceive crime statistics as measures of police performance. For example, when asked “What does your manager monitor or measure to assess whether you or your immediate team is performing well?” one first-line supervisor in a district responded: “Statistics; crime figures; files on hand” [QD041]. Crime statistics are monitored in the OPR frameworks and routine performance reporting of both agencies.

Few interviewees provided examples of managers reviewing quantitative measures of inputs (demand volumes), outputs (decisions), and outcomes (results) together to determine how they might be related. Neither was much evidence found of outcome performance measures being used to influence decisions that affect service levels. Almost all of the few examples that were found were given by PCC managers in New Zealand Police. For example, one PCC manager described attempting to identify differences in practices by comparing volumes of traffic tickets issued by different police personnel [NP076]. This same manager compared measures of police activity with numbers of road crashes to determine whether or not the activity is preventing crashes. Although quantitative performance information exists outside PCCs in New Zealand, I found negligible evidence of this information influencing decisions of police managers.

I did find one notable exception: A senior sergeant in charge of a police station [QD051] described using quantitative performance information to identify problems, evaluate impacts of actions, motivate personnel to undertake previously unpopular tasks, and
increase productivity of personnel. For example, this interviewee stated: “Recently I did a review on calls for service for 313 and 315 job codes that relate to disturbances, to monitor if a couple of things that had happened in the city have reduced the number of disturbance calls we’re going to which were a higher risk type of call”. To make well informed decisions, this interviewee proactively obtains performance information beyond the organisation’s routine performance reporting: “I’ve got my own download of CAD data for a month for particular issues or places or people. I review that CAD data on a weekly basis at least.” Additionally, this interviewee tasks staff with collecting information in a structured way, and rewards individuals and teams for collecting data. When asked if staff collect this information reliably, the interviewee responded: “Initially, no, but as the positives went out and the people were reminded, …and the confidence was built, and people were rewarded for doing a good job and filling it out, it’s very reliable now.”

When asked what led to this use of quantitative performance information, this manager [QD051] explained that this behaviour was initially triggered by community meetings that identified a systemic police performance problem in the local community. This caused the interviewee to review statistical data from police systems, which confirmed the existence of the problem. This insight caused the manager to subsequently make greater use of quantitative information. This manager’s practice developed out of a performance failure. This failure stimulated the manager to try a new approach to monitoring police performance. The competence that subsequently developed was self-initiated, rather than formally trained.
No other interviewee gave evidence of a similar level of effectiveness in using quantitative performance information to improve police effectiveness. However, some interviewees look at quantitative measures of demand and performance when deciding how to deploy staff. For example, when asked how performance monitoring helps, one first-line supervisor on a district responded:

Well it shows basically, I suppose, our clear-up rates and our calls for service, how we’re responding to them, where we need to improve things like directing personnel and resources. Where we can, we analyse the information and then work out our twenty day action plan around that, so that if we’re identifying we’ve got a problem on Friday night—alcohol related incidences at railway station, saying it’s an ongoing thing—then I can actually put an operational order together and direct staff to a particular area between certain times to try and combat our problems. [QD024]

However, when prompted to identify what this interviewee monitors to determine whether the performance of staff is good, the interviewee did not describe using quantitative information, but instead responded:

The quality of their logs; the information they put in their logs; verbal feedback in relation to a job they may have attended. When they come back to the station I can physically speak to them about it and say ‘what did you do about this?’ [QD024]

Although many interviewees are either sceptical or critical of quantitative performance information, some do make limited use of quantitative information outside the
organisation’s standard performance measurement frameworks. For example, one PCC
manager described using quantitative information to identify problems with employee
morale or commitment. This interviewee described using a formula developed by a
university to identify suspicious patterns in sick leave taken by staff: “If a person takes
say six weeks off sick leave they’ll hardly even appear on the graph, but if they take six
individual days in two months they’ll have a huge score. So, it’s trigger points for
interventions [sic]” [NP100].

In summary, although both agencies routinely report quantitative performance
information, I found few examples of such information being used in a consistent way
across an agency to influence decisions. The interviews provide some evidence of PCCs
monitoring demand volumes, call answer times, and response times. However, in
districts, examples of managers using quantitative performance information are isolated
and rare. If police decisions are rarely influenced by quantitative performance
information, it follows that routine performance measurement will have negligible
influence on police effectiveness.

5.3 RQ 2 Does Performance Measurement Help or Hinder Police
Effectiveness in Either Technical-rational or Institutional Terms?

Previous research presents mixed findings on the use and impact of performance
measurement in police (Ashby et al., 2007; Bayley, 2008; Hoque et al., 2004; M. H.
Moore & Braga, 2003; Savage, 2007). Benefits exist in police measuring their
performance (Hoque et al., 2004; M. H. Moore & Braga, 2003). However, it may be that
police organisations implement performance measurement frameworks to gain
organisational legitimacy by appearing progressive, rather than to improve service
delivery by improving their management of processes and resources (Willis et al., 2007). In this section I examine the influence of performance measurement on police effectiveness in both technical-rational and institutional terms by examining the attitudes to and behaviours surrounding performance measurement frameworks in police agencies. To answer the research question I analysed transcripts to address four more specific sub-questions:

**RQ2.1 What do police personnel believe is the reason for police measuring their performance in responding to calls for service?**

Understanding why police personnel in Queensland and New Zealand believe police measure their performance will enable assessment of the level of congruence between the criteria for police effectiveness identified in the literature and the perspectives of police personnel who deliver the service in Queensland and New Zealand. The literature reviewed in Chapter 2 identified numerous reasons why organisations would want to measure their performance. First, measuring performance informs management decisions that support effective and efficient delivery of services (Hoque et al., 2004; R. W. Scott, 1987). Second, performance measurement enables members of an organisation to make sense of what is happening in complex organisational contexts (Gill, 2011). Third, organisations that embrace NPM reforms—such as performance measurement—strongly are better able to cope with uncertainty in the environments in which they operate (Budding, 2004).

However, the literature also suggests that police agencies may use performance measurement to appear legitimate, rather than to improve their effectiveness. Public sector organisations are more likely to use performance information for external reporting
than for internal improvements (Taylor, 2009). Meyer and Rowan (1977) argue that organisations use such external performance reporting as a legitimising structure that acts as a buffer to enable internal actions and decisions to vary in response to practical considerations (p. 357). Willis et al. (2007) suggest that highly institutional organisations, such as police, experience pressure to adopt performance frameworks for structural conformity (isomorphism) with other organisations. DiMaggio and Powell (1983) explain this pressure as existing in organisations that operate in environments with a lot of uncertainty. These scholars argue that such organisations model themselves on other organisations to enhance their legitimacy and survive. In Section 5.3.1 I examine evidence from the interviews on the perspectives that police personnel in New Zealand and Queensland have on why police measure their performance.

RQ2.2 Do police personnel perceive performance measurement in police to be helpful?

The literature reviewed in Chapters 2 and 3 gives mixed results on whether performance measurement improves or reduces police effectiveness (Ashby et al., 2007; Bayley, 2008; Hoque et al., 2004; M. H. Moore & Braga, 2003; Savage, 2007). Results presented in Section 5.3.1 provide insights into how performance measurement might influence police effectiveness. However, this understanding is insufficient to determine whether performance measurement actually improves or harms police effectiveness. In Section 5.3.2 I examine evidence from the interviews on the impact of performance measurement on police effectiveness in responding to calls for service.

RQ2.3 What performance measurement mechanisms do police use?
In Chapter 2 I discussed Compstat and other operational/organisational performance review (OPR\textsuperscript{9}) mechanisms used by police agencies in USA and Europe. However, even though the Australian and particularly New Zealand public sectors strongly adopted NPM reforms in the 1980s and 90s, there has been little recent research published on how NPM has impacted on and evolved in Australian and New Zealand police agencies. The interviews sought to identify performance measurement frameworks used in these agencies. Results are presented in Section 5.3.3.

**RQ2.4 What performance measures do police monitor?**

As discussed in Chapters 2 and 3, some scholars (Legrand & Bronitt, 2012; Maguire, 2003; M. H. Moore & Braga, 2003) believe that performance measures monitored by police focus narrowly on measures that reinforce a police identity as society’s crime fighters, rather than reflect the majority of calls for service that citizens make. So, as well as exploring the different mechanisms that police use to measure performance, the interviews sought to identify performance measures monitored by New Zealand Police and Queensland Police Service. Results are presented in Section 5.3.4.

**5.3.1 RQ2.1 What do police personnel believe is the reason for police measuring their performance in responding to calls for service?**

Interviewees gave widely varying responses for the reasons why their agency measures its performance. These responses contained no dominant theme or pattern; no specific

\textsuperscript{9} In Queensland Police Service OPR stands for operational performance review; in New Zealand Police OPR stands for organisational performance review. I describe in this thesis how OPRs operated at the time I undertook Study 1. Both New Zealand Police and Queensland Police Service subsequently ceased to operate OPRs in the format they operated in 2009.
reason was identified by more than one third of interviewees. The strongest theme in the responses (n=38 out of 121 interviews) is that the agency measures its performance to help ensure appropriate resources are deployed; specifically, to help justify requests for additional resources or to plan where to deploy resources, as the following three responses illustrate: “For efficiencies in terms where we’ve got our resources and ah… you know, why are they there, and why would you have a certain number of resources in one place when they are getting more work in another” [QS047]; “Are we putting [our resources] in the right place? Do we need to put more money? If we’re asking for more money how do we justify that?” [NP079]; “The whole idea I suppose, is to measure; is to give the management groups an idea what is coming in so they can build their rosters, their deployment, their approach to tackling whatever problems that are within Districts” [NP121].

Some interviewees believed that senior managers measure police performance to “make themselves look good” [QD017]. Some interviewees associate this motivation with justifying the number of police personnel. For example, when asked what senior executives wish to find out from performance reports, one general duties officer stated: “Ways to make themselves look good [and] justify the numbers of police per community” [QG027]. This response provides an example explained by institutional theory of an organisation seeking to enhance its legitimacy to improve its prospects for obtaining resources and surviving (Meyer & Rowan, 1977; R. W. Scott, 2004; Suchman & Edelman, 1996).

Interviews also provide evidence of police using performance measurement to maintain their legitimacy following public criticism. For example, one PCC manager stated:
“When we’ve received some bad press around a complaint, we’ll try and get the message across of where we are now in terms of our service delivery” [NP073]. Similarly, one senior sergeant in charge of a police station described the implementation of performance measurement frameworks as “a process the organisation went through to try to restore public confidence” [QD037] following a crisis of legitimacy. A senior executive gave an example of this use of performance measurement by recounting an occasion when police response-time performance to burglaries had been declining. In order for performance to appear to be better, this senior executive changed how the response-time performance measure was calculated [NS060]. Similarly, a PCC manager [NP079] described how, following a particularly high profile crisis of legitimacy, police had a strong focus on reporting statistical measures of performance. However, as time passed and the crisis passed, emphasis on these performance measures reduced.

These responses provide examples of police personnel recognising that police are publicly accountable for their actions. Approximately one fifth (n=26) of interviewees stated that their agency measures its performance in responding to calls for service because police need to be publicly accountable, and measuring performance is a way of providing that accountability. Some interviewees expressed accountability as an obligation to the government because the government funds police. For example, one district manager stated: “I suppose managing performance at the end of it all comes back to being accountable to government for our funding is the overall performance measure at the end of the day [sic].” [QD036].

Approximately one fifth (n=24) of interviewees stated that their agency measures its performance to improve service delivery to the public, as the following responses
illustrate: “In order to in order to provide a better service” [QC005]; “To better, serve the public I suppose, and deal with stuff” [QG022]; “To improve service” [NP073]; “So that we can constantly look to improve our performance and provide better service to the public” [ND088]. No other purposes for measuring performance were identified by more than one fifth of interviewees. However, five further purposes were identified by at least 10 interviewees. These are: to ensure community expectations are being met (n=19), to ensure standards are being achieved (n=19), to help justify the actions of police (n=12), to improve efficiency (n=12), and to reassure the public (n=10). This result suggests that police personnel vary in their understanding of why police measure their performance. No consistent overarching purpose for police measuring performance is evident in the narratives of interviewees.

Nine interviewees did not know why their agency measured performance in responding to calls for service. Eight of these nine interviewees were from Queensland Police Service. Five of the eight were more junior in rank than sergeant. The most senior interviewee from Queensland Police Service who did not know why their agency measured performance was an inspector in charge of a division in a district. When asked why police measure their performance in responding to calls for service, this inspector responded: “Well I wouldn’t know, I know we don’t, it’s a measure that I have, ‘cause I like, it gives um public confidence that we have resources to respond to people’s needs” [QD056]. Although only one New Zealand Police interviewee stated he did not know why his agency measured performance, this interviewee held a senior position as a superintendent in charge of a district. Although this superintendent did not know why his agency measured performance, he did not criticize the performance measurement frameworks in operation. Instead, he seemed to trust that they had a purpose, and he
expressed a hope that performance measurement was used to help achieve the organisation’s strategic plan: “I’d like to sort of think that in a broad and general sense that the performance measures are really about trying to meet, say as an example, the six [strategic priority] outcomes that are in place” [ND072].

In summary, responses to the question of why police measure their performance (Question 6) reveal that police personnel vary in their understandings of the purpose of performance measurement in police. Many interviewees perceive that police measure their performance to achieve technical-rational objectives such as planning deployment of resources, being accountable to the government and/or the public, or improving service delivery. However, many responses provide tangible examples of police measuring performance to achieve the institutional objective of maintaining legitimacy. In particular, interviewees gave examples of police using performance measurement as a tool to counter public criticism. This finding is consistent with those of (Taylor, 2009) and (Willis et al., 2007), and support the explanation provided by institutional theory (Meyer & Rowan, 1977; R. W. Scott, 2004; Suchman & Edelman, 1996) that institutionalised organisations strive to maintain their legitimacy.

5.3.2 RQ2.2 Do police personnel perceive performance measurement in police to be helpful?

In this section I describe findings about the perspectives that police personnel have of the influence of performance measurement on police effectiveness. Fifty seven percent (n=69) of interviewees perceived that performance measurement is helpful. Twenty seven percent (n=33) of interviewees were either unsure or perceived it is sometimes helpful and sometimes unhelpful. Sixteen percent (n=19) of interviewees perceived that
performance measurement is unhelpful. Interviewees from New Zealand Police expressed similar perspectives to those from Queensland Police Service.

Interviewees were asked how performance measurement helps to achieve the objectives they had identified earlier in the interview. Many (n=33) interviewees could not identify how this occurred. Most of the 19 interviewees who believed performance measurement is unhelpful either said this is because if performance measurement is effective they would have been allocated more staff, or said they had not observed any improvements over time as a result of performance measurement. For example, when asked to recall an instance of performance being assessed as ‘not good enough’, and what happened as a result of this assessment, a general duties officer responded: “Nothing; it just gets swept under the carpet. Um… it’s too-bad-so-sad type of thing. It… nobody’s… doesn’t seem to be accountable for that kind of stuff” [QG042]. Similarly, a senior sergeant in charge of a police station stated: “The issue is what they do with that if they have an underperforming manager. I don’t think they manage that situation very well” [QD037].

Some interviewees were cynical of performance measurement, and expressed a desire to be left alone to do their job as they saw fit. For example, when asked if the performance monitoring that police do is helpful, one general duties officer responded: “No; it’s obstructive when you’re trying to run a section of staff and you want them on the road” [NG092]. When asked what matters in police responding effectively to calls for service another general duties officer responded: “Allowing police to do what police do” [QG027]. Similarly, a communications room operator stated: “The amount of police surveys they do drive me crazy because I don’t know if anything ever comes of them” [NC104]. Two performance managers said that many people within police see
performance measurement as a threat because they do not want to be held accountable. For example, one of these performance managers stated:

There is a reluctance by many in senior and middle management in police to have any type of performance reporting because there is a feeling that it holds them to account, and they would rather not be held to account, rather than looking at performance data as a way of measuring how we are doing and what we need to do better, or more or less of. [NM062]

Even though over half of all interviewees believed that performance measurement is helpful, few were able to substantiate this belief with examples of instances where police are influenced by performance measures to make changes that improve effectiveness. Similarly, most interviewees struggled to identify how performance measurement influences their own actions. For example, when asked for an example of how performance monitoring has affected or influenced the interviewee, one communications room operator responded: “No, I cannot” [QC005], and a first-line supervisor in a district responded:

If um… If we know of um… of they previous [sic] urgent ah… or violent person living at an address, I’ll know to ah… to advise and send more units… or a… or if there’s um… like for instance with the mother and the son, if ah… that happens again, we’ll… then we’ll know the son is violent, and we’ll know to get a car there more urgent [sic]. [QD041]

Some interviewees believed that performance measurement is useful because knowing that a measure is being monitored makes managers stay focussed on the things that are
most important. Similarly, some interviewees believed performance measurement is useful because it creates awareness of accountability: “Knowing your performance is measured, everything you do is measured, means that you can’t get away with being slack, and you can’t get away with doing something wrong” [NC114]. However, other interviewees were concerned that some performance measures are counter-productive because they are poorly designed, creating focus on the wrong things. One performance manager did not know what police should be measuring: “One of the things—and my GM and I have discussed it—that we are not sure about is whether we are targeting the right things, and if so how do we know, and if so where’s the data to support that” [NM117]. This perspective accords with literature that argues that police struggle to measure the right measures (Brunetto & Farr-Wharton, 2005; Fleming, 2009; Lawton et al., 2000; Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; Reiss, 1992; Skogan & Frydl, 2004; Sparrow et al., 1990). However, overall, the findings presented in this section are inconclusive. The majority of interviewees perceive that performance measurement is helpful. However, few were able to give examples of how performance measurement had helped. Some interviewees believed that performance measurement does not help police to be effective, because police management does not take appropriate action when performance deficiencies are identified.

Study 1 is a qualitative study; it is unable to measure the actual impact of performance measurement on the effectiveness of police in responding to calls for service. To determine whether performance measurement does influence police effectiveness in responding to calls for service, I undertook a quantitative study described in the next chapter. That study—Study 2—examines the relationship between indicators of police effectiveness in responding to calls for service and performance measures monitored by
police. To form specific testable hypotheses, I need to identify appropriate performance measures and indicators of police effectiveness. The remainder of Section 5.3 focuses on identifying the performance measurement frameworks and measures used by police.

5.3.3 RQ2.3 What performance measurement mechanisms do police use?

Six different mechanisms for measuring or monitoring police performance or effectiveness were commonly referred to by interviewees. These include OPRs, routine performance reporting, individual performance appraisals, complaints, qualitative checking, and the public media. The proportions of interviewees from each police agency who referred to these mechanisms is presented in Table 5.1 Each of these six mechanisms will be described.

Table 5.1
Proportion of Interviewees who Referred to each mechanism

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPRs</td>
<td>11%</td>
<td>41%</td>
<td>26%</td>
</tr>
<tr>
<td>Routine performance reporting</td>
<td>10%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>Individual performance appraisals</td>
<td>33%</td>
<td>26%</td>
<td>30%</td>
</tr>
<tr>
<td>Complaints from citizens</td>
<td>49%</td>
<td>48%</td>
<td>49%</td>
</tr>
<tr>
<td>Checking individual instances</td>
<td>29%</td>
<td>45%</td>
<td>36%</td>
</tr>
<tr>
<td>Checking public media</td>
<td>21%</td>
<td>10%</td>
<td>16%</td>
</tr>
</tbody>
</table>

Over a quarter of interviewees mentioned monitoring measures of demand volumes or monitoring local performance measures. However, most interviewees did not indicate whether these measures were monitored routinely. The proportions presented in Table 5.1 for routine performance reporting are conservative, as they only include interviewees who explicitly referred to regular routine performance reporting.
**Mechanism 1: OPRs**

As discussed in Chapter 2, the New York Police Department’s Compstat performance framework spread throughout the world (with adaptations). Queensland’s version of Compstat is called operational performance review (OPR), and includes aspects of Compstat (Mazerolle et al., 2007). New Zealand Police operate a version of Compstat called organisational performance review (which I will also refer to as OPR). Both the Queensland and New Zealand version of OPR involves a formal internal performance review of a part of the organisation. The particular mechanisms and frequency of OPRs differ between the two agencies. However, they contain a number of common features: OPRs are facilitated by a centralised group from the agency’s headquarters, and involve personnel from the part of the organisation that is being reviewed; OPRs include monitoring *Key Performance Indicators* (KPIs) that remain relatively stable between successive reviews, making it possible to assess whether performance is improving or deteriorating; OPRs are applied at the level of a large work group such as a district that is the responsibility of a superintendent; OPRs have the effect of generating a cascading monitoring framework within districts in which similar KPIs are monitored at a more localised level.

As shown in Table 5.2, the narratives about OPRs of interviewees varied between the two agencies and between different role-types. OPRs were more evident in Queensland Police Service than in New Zealand Police, with 41% of Queensland interviewees referring to OPRs, compared to only 11% of New Zealand interviewees. This difference was most evident amongst PCC personnel: communications room operators and PCC managers. No New Zealand Police PCC personnel mentioned OPRs, and no interviewee from New Zealand Police mentioned an OPR ever having occurred in a PCC. However, some
Queensland PCC personnel referred to at least one OPR having previously occurred at their PCC, and approximately one third of PCC personnel in Queensland referred to OPRs. Queensland Police Service interviewees also generally provided more detailed descriptions or more examples of personally being involved in some way with the OPR mechanism than did New Zealand Police interviewees. For example, one communications room operator stated: “The last OPR we did was probably two or three years ago, I did graphs on calls, abandoned calls, staffing, total jobs, all that type of thing” [QC008].

Table 5.2
Proportion of Interviewees who Referred to OPRs

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications room operators</td>
<td>0%</td>
<td>25%</td>
<td>13%</td>
</tr>
<tr>
<td>General duties officers</td>
<td>8%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>PCC managers</td>
<td>0%</td>
<td>45%</td>
<td>23%</td>
</tr>
<tr>
<td>District managers</td>
<td>17%</td>
<td>59%</td>
<td>37%</td>
</tr>
<tr>
<td>Senior executives</td>
<td>25%</td>
<td>50%</td>
<td>38%</td>
</tr>
<tr>
<td>Performance managers</td>
<td>33%</td>
<td>100%</td>
<td>56%</td>
</tr>
<tr>
<td>All roles</td>
<td>13%</td>
<td>41%</td>
<td>26%</td>
</tr>
</tbody>
</table>

In both agencies, performance managers more frequently referred to OPRs than did any other role type. More junior groups—communications room operators and general duties officers—referred to OPRs less frequently than did any management group outside PCCs. It is unsurprising that performance managers—who were all based at state headquarters—more frequently referred to OPRs than did any other role type, because OPRs are facilitated by a centralised group at state headquarters.
Senior managers generally perceived OPRs as being important and useful in supporting police to be effective. In particular, some senior executives described OPRs as having created a culture of accountability for performance, rather than just accepting the status quo. For example, one senior executive described the OPR framework as “critical”, and said that OPRs had significantly changed the organisation. This interviewee stated: “Before OPR—it’s hard to imagine a time before OPR—but before OPR there was nothing. Nobody asked any questions. Nobody had any expectations of anybody. I can’t believe we had a time before it.” [QS047]

One performance manager believed that it was not what occurred in an OPR that created performance improvement; the act of announcing the intention to undertake an OPR stimulated the business group that was about to be reviewed to take action to improve the performance of the business group before the review occurred [QM007]. This view aligns with that of researchers who studied the impact on crime volumes of introducing OPRs to Queensland Police Service (Mazerolle et al., 2007). These scholars speculated “that the communication and extensive consultation about the introduction of the OPRs would have created what is known as an ‘announcement effect’ (Smith et al., 2002) such that District Commanders would have altered their management and operational strategies at this time” (Mazerolle et al., 2007, p. 242). The study by Mazerolle et al. (2007) reflected this perspective by using an interrupted time-series analysis methodology to measure changes in crime volumes from a fixed point in time that corresponded to announcement that an OPR would be undertaken. There is a gap in the literature on whether performance information subsequently monitored through OPRs influences performance outcomes such as crime volumes. This thesis seeks to address this gap by identifying performance
measures monitored in police performance frameworks such as OPRs, and by testing whether variations in these measures influence measures of police effectiveness.

**Mechanism 2: Routine performance reporting**

Some interviewees\(^{11}\) described a practice of routinely reporting performance information to their superiors. Ten mentioned a regular reporting cycle such as monthly or weekly. For example, one senior sergeant stated: “We have a district operations meeting every month and at those meetings I’m required to report to the district officer in relation to our response to our crime trends” [QD019]. However, others were less specific about the frequency or content of such performance information. Nevertheless, it was apparent from the interviews that performance measurement and reporting frameworks exist throughout all levels of management in both Queensland Police Service and New Zealand Police.

Some interviewees described one effect of OPRs as stimulating a cascading monitoring framework within districts. Similar KPIs are monitored on a frequent and regular basis at a more localised level. For example:

OPRs are done at a number of levels. The region does an OPR with the commissioner, but in a sense the officer-in-charge meetings are a mini OPR whereby they throw all your stats up and you talk to your stats. If there’s a spike in an area they’ll want to know why the spike is happening, and you’re going to need to know why. They take that information, and then when they go to their regional one where the district goes to the commissioner, then they’ll do the same thing, and they’ll feed a lot of that information up that you’ve told them. And then

\(^{11}\) See footnote 10 under Table 5.1
when the region goes to the commissioner for the regional OPR they’ll do exactly the same thing. [QD037]

Although routine performance reporting often reports similar measures to the measures used in OPRs, routine performance reporting within districts and PCCs can be considered a distinct mechanism. Unlike OPRs, routine performance reporting is often initiated and operated within a district or PCC rather than by a centralised group at headquarters, and reporting occurs more frequently than OPRs. For example, one PCC manager described various routine cycles of reporting to their manager, including “…the weekly report, and then there’s the monthly report” [NP118].

**Mechanism 3: Individual performance appraisals**

Both Queensland Police Service and New Zealand Police operate an individual performance appraisal framework. In this framework all managers and supervisors are expected to formally review, every six or twelve months, the performance of each employee who directly reports to them. Performance is reviewed against criteria set in advance. These criteria usually relate to the core functions of the employee’s position.

As shown in Table 5.3, 30% of interviewees identified individual performance appraisals as a performance measurement mechanism. Senior executives and performance managers made little mention of individual performance appraisals, but numerous district managers (40%), PCC managers (36%), communications room operators (33%), and general duties officers (22%) did discuss individual performance appraisals. Interviewees varied widely in their level of engagement with the individual performance appraisal framework. Some interviewees stated that they had not experienced a performance appraisal for a long time. For others, appraisals occur regularly. Some interviewees saw performance appraisals as
purely compliance exercises or were ambivalent about them. Others believed performance appraisals were useful, at least to force a regular check that would otherwise be neglected. For some interviewees, performance appraisals are the only mechanism by which they receive any feedback on their performance.

Table 5.3
Proportion of Interviewees who Referred to Individual Performance Appraisals

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications room operators</td>
<td>50%</td>
<td>17%</td>
<td>33%</td>
</tr>
<tr>
<td>General duties officers</td>
<td>25%</td>
<td>18%</td>
<td>22%</td>
</tr>
<tr>
<td>PCC managers</td>
<td>36%</td>
<td>36%</td>
<td>36%</td>
</tr>
<tr>
<td>District managers</td>
<td>39%</td>
<td>41%</td>
<td>40%</td>
</tr>
<tr>
<td>Senior executives</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Performance managers</td>
<td>17%</td>
<td>0%</td>
<td>11%</td>
</tr>
<tr>
<td>All roles</td>
<td>33%</td>
<td>26%</td>
<td>30%</td>
</tr>
</tbody>
</table>

Informal feedback by one’s supervisor outside of the formal performance appraisal interviews also occurs in both agencies. The extent of such feedback varies greatly. Some interviewees meet at least once a day with their supervisor, and feedback on performance is frequently given. Others meet rarely with their supervisor, and performance is rarely discussed.

Several interviewees, from various ranks and roles in both agencies, stated that they only receive feedback on their performance when such feedback is negative. Most interviewees interpret lack of feedback as indicating that they are performing satisfactorily. One interviewee stated that the way they know whether or not things are
working well in police responding to calls for service is by “whether or not we are being yelled at by the bosses” [QC001].

**Mechanism 4: Complaints from citizens**

Almost half of all interviewees (56 of 121) did not identify any of the proactive structured performance measurement mechanisms of OPRs, routine performance reporting, or individual performance appraisals. However, more than two thirds of interviewees described reactive qualitative mechanisms for assessing whether police are performing well in delivering a service of responding to calls for service. One superintendent in charge of a district believed police do not “have good systems to measure how well it is done” [ND095]. When asked why, this interviewee stated: “Probably I guess, history. We haven’t in the past and we’ve relied on complaints to make some pretty non-specific assessment of whether or not it’s working.”

More interviewees (49%) identified complaints from citizens as a mechanism used to monitor the police performance than identified any other mechanism. Except for the three performance managers in Queensland Police Service, this mechanism was referred to throughout all levels and roles in both agencies. However, the formal complaints frameworks of both agencies are primarily designed and used as disciplinary investigation frameworks, rather than as customer feedback mechanisms. One exception to this was found in New Zealand Police PCCs, which maintain a ‘Praise and Complaints Database’ that is used for recording customer feedback. However, New Zealand Police mainly use this database to record and manage actions in response to each record of praise or complaint, rather than to monitor and analyse patterns in performance measures. Both agencies have extensive organisational policy and procedural structures for managing
complaints, both internally and for working with the external police complaints authority\textsuperscript{12}.

Many managers said that they monitor complaints as a way of knowing whether or not things are working well in police responding to calls for service. For example, when asked what is most helpful to monitor or measure in police responding to calls for service, one first-line supervisor in a district responded: “Well I suppose if we’re not getting any complaints then we’re reasonably satisfied that we must doing a fairly good job” [QD038]. However, few interviewees count or analyse the number and nature of complaints in any systematic way. Managers interviewed commonly stated that responses to complaints usually seek to establish whether the police action that led to the complaint could be justified. For example, one PCC manager stated: “Depending on what the issue is, if it’s simply a performance issue or something that we can explain away, it’s usually just written up and explained away about that, and we would enter that in our complaint data bases as probably an expression of dissatisfaction with our service” [NP073].

**Mechanism 5: Monitoring the quality of work in individual instances**

Although complaints is the most frequently mentioned mechanism for monitoring performance, when discussing how performance is actually managed, interviewees more readily described qualitative checking of individual jobs by an individual’s supervisor and by the supervisor’s managers. With general duties officers, this involves supervisors and their managers checking the reports written by the officer who attended the job. For example, one general duties officer stated: “We just put in a report, and then it goes to the

\textsuperscript{12} In New Zealand the external police complaints authority is the Independent Police Conduct Authority; in Queensland at the time Study 1 occurred the equivalent organisation was called the Crime and Misconduct Commission. It is now been renamed as the Crime and Corruption Commission.
sergeant, and he has a look at it, and then someone else has a look at it, and then if something’s wrong with it it’ll come back to the first person that responded to it” [QG018].

Supervisors check that reports are complete and well written, that actions taken comply with standard operating procedures, and whether any actions could have been done better. Where deficiencies are identified these are communicated to the officer concerned. For example, one PCC manager stated: “If I’m reading every job there might be three or four jobs in a day—probably not a lot more than that—that I’ll look at and think ‘I’m not happy with that action’, so I’ll go over to the dispatcher and say ‘Hey listen’” [NP101]. Similarly, an inspector in charge of a division in a district stated:

I look at the occurrence sheets for the three stations under my command. I go through what the troops do every day just to see whether what they put into the occurrence sheets is relevant or not, but to make sure what they’ve done is appropriate for the job. So, if they’re going to the break and enter they’re doing what they should be doing. And then occasionally I’ll pull out the crime reports and make sure that the quality’s there. And from time to time I’ll send emails to various officers and say “look, this particular job here, you know, it’s a very serious matter, but you’ve been rather brief in your report.” [QD040]

**Mechanism 6: Public media**

Managers—particularly senior managers—monitor commentary appearing in the media or parliament that reflects negatively on police. They consider the amount of noise in the media to be a useful and/or important indicator of police performance. Interviewees often mentioned noise in the media and noise in parliament together as being similar indicators,
because stories in the media often stimulate political debate in parliament, as illustrated in a statement by one senior PCC manager:

Generally we find that we have an increase in ministerial or media requests when the service delivery times are extended. In other words, usually if we see an increase in the number of complaints coming from the public because police didn’t answer triple-0\textsuperscript{13} in a short time frame or didn’t get to the job in a timely manner, and usually if we don’t answer calls within that 30 seconds for triple-0, [or the delay in answering] goes to a minute or two minutes, [we can] almost guarantee we’ll end up with a ministerial or media [enquiry]. Or if the police don’t get to a code one or two job within five to ten minutes, automatically we’ll get a complaint. Or even if it is a minor, what say, if it takes more than forty minutes, we just know we’re going to get a media [enquiry] or a thing through the minister’s office.” [QP044]

Interviewees from both agencies indicated that their agency is sensitive to organisational risk, in particular, public criticism of police performance or integrity. One performance manager perceived that: “Because of the nature of policing, things become far more political than in some cases they deserve to be” [QM011]. Interviewees more readily perceived the public media as seeking to criticise police, rather than seeking to support police or be objective. For example, when asked about the purpose that the media has in making requests under the Official Information Act for information about police response to calls for service, a superintended in charge of a district responded: “Really they are looking for service failures in my view” [ND089]. Similarly, a senior sergeant in charge

\textsuperscript{13} Triple-0 (000) is the emergency phone number used in Australia. This is equivalent to 111 in New Zealand, 999 in Britain, and 911 in USA.
of a police station stated: “If it’s in the paper it’s generally criticism of the police” [ND090]. One senior executive stated that when police fail to provide service “in a timely or appropriate manner, you get complaints through to the media.” [QS046]

Some interviewees value the public media as a source of information on how effective police are. For example, one general duties officer stated: “It’s amazing how much we work out through the media as to how effective things are” [QG039]. Many senior managers routinely monitor media stories about police, and are interested in parliamentary debates about police performance. However, no evidence was found of any manager systematically counting or analysing over time the volume of, or patterns in, media stories or parliamentary debates about police performance or effectiveness. The most strongly articulated reason managers gave for the interest they have in media stories is the importance of police retaining public trust and confidence. For example, one PCC manager stated: “The bad headline can have a huge impact on the organisation, or even on the government itself” [NP100].

Some interviewees view complaints and media stories as similar types of performance indicators; both are usually reactive reports of individual instances of failures. For example, when asked what indicated whether things were working well with police responding to calls for service, one senior sergeant in charge of a police station responded: “Because I’ll see it in the media the next day—full stop!” [QD023]. Similarly, when asked what is most helpful to monitor or measure, a superintendent in charge of a district responded:

Well, it’s not the most important, but it’s the thing I do every day, and probably every district commander, every area commander does it, just opens the paper and
has a look, and then there’s an assessment of ‘is it an issue or isn’t it’. Most of the time you’ll read it and think: ‘Well, that’s a beat-up you know’; that’s not really an issue. But occasionally you might open the paper and say: ‘Well this is the third thing I’ve read in the last week about this’; you know—a bit of an issue. [ND089]

In addition to the six performance measurement mechanisms discussed in this section, some interviewees suggested that public satisfaction surveys are a means by which police performance could be measured. However, none gave any evidence of survey results actually influencing decisions. This may be because survey results are published too late to be valued by police. One senior executive [QS047] believed that satisfaction surveys have little influence on decisions because results take many months to be published. This perspective is consistent with the finding by Colvin and Goh (2005) discussed in Chapter 2 that police attitudes are strongly influenced by timeliness.

5.3.4 RQ2.4 What performance measures do police monitor?

A range of measures were identified in the interviews. These include response time, complaints, offence volume, crime resolution rate, volume of calls for service, and volume of repeat calls for service to the same address. The proportions of interviewees from each police agency who referred to these mechanisms is presented in Table 5.4. Each of these six measures will be described. Some interviewees also identified measures of customer satisfaction, dispatch timeliness, and/or efficiency of handling phone calls as measures of police performance in responding to calls for service. However, none of these three measures were identified by more than 17 percent of interviewees.
Table 5.4
Proportion of Interviewees who Referred to each measure

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response time</td>
<td>63%</td>
<td>48%</td>
<td>56%</td>
</tr>
<tr>
<td>Complaints</td>
<td>49%</td>
<td>48%</td>
<td>49%</td>
</tr>
<tr>
<td>Offence volume</td>
<td>27%</td>
<td>33%</td>
<td>30%</td>
</tr>
<tr>
<td>Crime resolution volume/rate</td>
<td>19%</td>
<td>29%</td>
<td>24%</td>
</tr>
<tr>
<td>Repeat calls for service</td>
<td>2%</td>
<td>29%</td>
<td>15%</td>
</tr>
<tr>
<td>Calls for service volume</td>
<td>30%</td>
<td>33%</td>
<td>31%</td>
</tr>
</tbody>
</table>

**Response time**

Literature reviewed in Chapter 3 (Bayley, 1994b; Eckblom & Heal, 1982; Herbert, 2006; Silberman, 1978) identified the timeliness of police response to calls for service as being an important dimension of police effectiveness. Consistent with this perspective, over half of all interviewees identified response time as a performance measure (see Table 5.5). More interviewees in New Zealand Police (63%) than in Queensland Police Service (48%) identified response time as a performance measure. Apart from the smallest group—performance managers in Queensland Police Service (n=3)—at least half of all management groups in both agencies identified response time as a performance measure.
Table 5.5
Proportion of Interviewees who Identified Response Time as a Measure

<table>
<thead>
<tr>
<th>Role</th>
<th>New Zealand</th>
<th>Queensland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications room operators</td>
<td>50%</td>
<td>25%</td>
<td>38%</td>
</tr>
<tr>
<td>General duties officers</td>
<td>42%</td>
<td>36%</td>
<td>39%</td>
</tr>
<tr>
<td>PCC managers</td>
<td>64%</td>
<td>82%</td>
<td>73%</td>
</tr>
<tr>
<td>District managers</td>
<td>78%</td>
<td>53%</td>
<td>66%</td>
</tr>
<tr>
<td>Senior executives</td>
<td>75%</td>
<td>50%</td>
<td>63%</td>
</tr>
<tr>
<td>Performance managers</td>
<td>83%</td>
<td>33%</td>
<td>67%</td>
</tr>
<tr>
<td>All roles</td>
<td>63%</td>
<td>48%</td>
<td>56%</td>
</tr>
</tbody>
</table>

**Complaints**

The interviews explored how the interviewee knows whether things are working well in police responding to calls for service, and how the interviewee’s agency measures or manages performance in responding to calls for service. When asked about measures of police effectiveness or performance in responding to calls for service, just under half of the interviewees in both agencies mentioned complaints against police. For example, an inspector in a district stated: “Generally speaking if you’re having problems in calls for service you’ll get a large number of complaints from the community” [QD017]. A senior sergeant in charge of a police station stated: “If the complaints are down—if the crime’s down, and that sort of thing—well then we must be going okay” [QD028]. Similarly, when asked: “How do you know whether things are working well or not in police responding to calls for service from the public?” a sergeant in charge of a team of general duties officers responded: “Lack of complaints from the public” [ND069]. However, although interviewees frequently referred to volumes of complaints as an indicator of
police performance or effectiveness, few actually count or monitor trends in numbers of complaints.

Almost half of all interviewees in each agency perceived that their agency measures volumes of complaints, but when prompted, all struggled to provide evidence of numbers of complaints being counted or reported. For example, when asked “How do you know whether things are working well or not in police responding to calls for service from the public?” one senior executive described looking at “the type of incidents that we’re dealing with, and complaints that we’re getting from the public, and those have decreases” [NS061]. However, when the interviewer prompted: “I’m wondering how you count the number of complaints”, the interviewee responded: “Well for me I only deal with the high profile ones”. Similarly, when asked: “Do you count the number of complaints coming through?” a senior sergeant in charge of a police station responded: “Not in a formalized manner. We do keep a—well no—I suppose we do actually, because we keep a record of them here” [ND086]. When asked if complaints are counted, one senior executive responded: “No; no we don’t” [QS045]. Only one interviewee described reporting or monitoring patterns in volumes of complaints [NS047].

Rather than monitor quantitative information about complaints as part of any performance measurement framework, interviewees more commonly discussed activity in response to a complaint aimed at reinforcing the organisational legitimacy of police. For example, one PCC manager stated: “When we’ve received some bad press around a complaint, we’ll try and get the message across of where we are now in terms of our service delivery” [NP073]. This manager explained that responses to complaints usually seek to establish whether the police action that led to the complaint could be justified: “Depending on what
the issue is, if it’s simply a performance issue or something that we can explain away, it’s usually just written up and explained away about that” [NP073].

Whether or not complaint volumes are quantified and monitored, over half of the interviewees in the PCC managers, district managers, and senior executives groups identified complaints as a performance measure or performance measurement mechanism. Senior managers, in particular, readily identified complaints (75% of superintendents and 88% of senior executives). As shown in Section 5.3.3, senior managers also monitor commentary appearing in the media or parliament that reflects negatively on police. They consider the amount of noise in the media to be a useful and/or important indicator of police performance. For example, one senior executive stated:

I’m not a great one for gathering data for the purpose of having data on something. I think that often times you can actually get a feel for what’s going on out there just by gathering information from a whole range of sources whether it be the media, local politicians, or whatever. So, knowing that there’s, sort of, you know, ten complaints as opposed to twelve or fifteen I don’t think is a useful measure. [QS045]

Crime and resolution volumes

As discussed in Chapters 2 and 3, prior research found that police monitor performance measures related to crime and resolving crime or arresting offenders (Legrand & Bronitt, 2012; M. H. Moore & Braga, 2003). Consistent with this prior research, interviewees in Study 1 identified measures related to crime and resolving crime as performance measures. Thirty percent of interviewees identified the volume of offences as being a performance measure. District managers most frequently identified this measure (49%);
communications room operators (21%) and general duties officers (22%) least frequently identified this measure. Twenty four percent of interviewees identified crime resolutions (apprehension of offenders) or resolution rate as a performance measure.

Some scholars (Maguire, 2003; Reiss, 1992; Sparrow et al., 1990) previously suggested police focus narrowly on measures that reinforce a police identity as society’s crime fighters, rather than reflecting the majority of calls for service that citizens make. Results of Study 1 provide only limited support for this view. One senior PCC manager stated:

Most of our OPR’s are focused on crime trends and crime patterns, and relate to property crime, personal crime, some emerging events and major incidents. It [sic] also focuses on finances, HR and complaints, but there is no service standard for triple-0 calls or for service delivery—still isn’t to this day. So, most police are more focused on the crime management model than they are on the service delivery. Is it important? I think it is, but at the moment they don’t really have the data to spend a lot on it. Are bosses keen to make sure they get to the jobs? Yes, because they don’t want a complaint so they’re being driven by their complaints data more so than a service delivery model. [QP044]

Although offence volumes and resolution rates are monitored by both Queensland Police Service and New Zealand Police, interviewees also identified a broader range of performance measures. Response time (56% of interviewees), complaints (49%), and demand volumes (31%) were identified more frequently as performance measures than were offence volumes (30%) and resolution rates (24%).
**Repeat calls for service**

A focus on reducing repeat calls for service to the same address was strongly evident in interviews with district managers in Queensland Police Service. OPRs and Routine performance reporting within districts in Queensland Police Service include monitoring repeat calls for service to the same address. Eleven of the 17 Queensland Police Service district managers who were interviewed identified repeat calls for service as a performance measure. Six other interviewees from Queensland Police Service, but only one interviewee from New Zealand Police\(^\text{14}\), identified repeat calls for service as a performance measure.

Interviewees view repeat calls for service as an indicator of police effectiveness and efficiency: “It’s about attending to calls for service in a way that they don’t reoccur, like addressing a problem” [QD051]. Some interviewees consider a repeat call for service to be a waste of police time because a repeat call for service can be avoided if the quality of the initial response is good. For example, one senior PCC manager stated that police focus on the quality of the police response to domestic violence, to ensure “that matter is dealt with appropriately [and] that it doesn’t then lead to repeat calls for service that are going to take time and energy” [QP034]. One senior sergeant in charge of a police station described the purpose of monitoring repeat calls for service as being “to identify repeat calls for service locations, and then strategies that we can put in place to reduce calls for service, because obviously by reducing calls for service and reducing reported crime and, that’s our core business [sic]” [QD019]. A superintendent in charge of a district described the role of an intelligence analyst attached to a PCC: “Their job is to do an analysis of our

\(^{14}\) When the Study 1 interviews took place, New Zealand Police did not routinely measure repeat calls for service.
repeat calls for service to look at what steps can we take engaging with other agencies or non-government agencies or organisations to actually cut down on the number of calls for service” [QD043].

**Calls for service**

The performance measures identified by interviewees include a mix of inputs (demand volumes), outputs (management decisions), and outcomes (results). Interviewees did not generally distinguish between these concepts in the interviews, even though only outcomes reflect police effectiveness. For example, demand volumes may indicate how busy police are, but they do not indicate whether police are effective. However, the timeliness of police response is a measure of police effectiveness; it affects how satisfied the caller is with the service received from police, and it may also affect the likelihood of preventing harm or apprehending offenders (Eckblom & Heal, 1982; Parks, 1984; T. Pate, 1976; Percy, 1980).

Just under a third of all interviewees referred to demand volumes, such as the volume of calls for service received by police, as being performance measures. PCC managers most frequently (55%) identified calls for service as a performance measure. Interviewees described patterns in demand volumes as being useful “to justify a request for more resources” [QM011] or “to determine if we have enough resources at a particular time of the day or night” [QD023]. Many managers were unable to describe how performance measurement influences their actions.

If managers are not influenced by performance measurement, employees may perceive there to be a conflict between formal organisational goals as expressed through performance measures and the behaviours of managers. Literature reviewed in Chapter 2
argues that loose coupling exists where agents are disconnected from organisational goals (Weick, 1990). Perceptions by an organisation’s employees of misalignment or confusion between goals can lead to loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977). Loose coupling manifests as a gap between formal structures and work practices. Formal rules are often broken, and decisions are often not implemented (P. Smith, 1993). It follows that loose coupling, if present, may affect police effectiveness.

5.4 RQ3 What Influence Does Performance Measurement in Police Organisations Have on Loose Coupling, and What Influence Does Loose Coupling Have on Police Effectiveness?

Literature reviewed in Chapter 2 identifies conditions that can lead to loose coupling in an organisation. This literature also identifies how loose coupling manifests in an organisation. So, to answer RQ3 I first explore whether conditions that can lead to loose coupling are present in New Zealand Police and Queensland Police Service. In Section 5.4.1, I examine the perspectives that police personnel of different roles and ranks have about police effectiveness. This examination will reveal differences in perspectives of junior and senior personnel that are consistent with conditions that give rise to loose coupling. Then, in Section 5.4.2, I present results for evidence on the existence and nature of loose coupling in the two police agencies.

5.4.1 Police Perspectives on What Constitutes Effectiveness in Responding to Calls for Service

Literature reviewed in Section 3.3 identifies numerous dimensions of police effectiveness in responding to calls for service, but provides little evidence on whether police personnel
understand effectiveness in terms of these dimensions. Dimensions such as bringing 
offenders to justice (Silberman, 1978), preventing harm, and maintaining order (Bayley, 
1994b) directly contribute to maintaining a safe and ordered society. Dimensions such as 
demonstrating procedural justice (Tyler, 1990) and being responsive to citizens (Herbert, 
2006) contribute indirectly by influencing citizens to feel safe, respect the law, and 
support the police (G. Cordner, 1995; Goldstein, 1987a; Sparrow et al., 1990).

Analysis of the interview transcripts suggests that New Zealand and Queensland police 
personnel understand effectiveness in similar terms to the dimensions identified in the 
literature, but that some differences exist between the perspectives of junior and senior 
personnel. Five topics were identified in response to questions on what interviewees 
believe constitutes effectiveness in responding to calls for service. These topics are 
timeliness of response, apprehension of offenders, prevention of harm or restoration of 
order, attendance, and satisfaction of the caller. Each of these topics will be discussed.

**Timeliness of response**

Three-quarters of interviewees (n=89) identified the timeliness of police response as 
important. Results were similar for New Zealand and Queensland. At least half of the 
interviewees in each of the six role types in both police agencies identified timeliness of 
response as important. Some interviewees referred specifically to speed of police 
response to urgent matters where a threat to life or property exists; most were not specific 
about what type of calls are important for police to attend quickly. Interviewees gave a 
range of reasons for rapid response being important. Common reasons given include: to 
maximise the chance of apprehending offenders, to prevent harm, and to satisfy the caller. 
The following examples illustrate these reasons:
In describing the outcome of a particular call for service a first-line supervisor in a PCC stated “I would call it successful because we contained the situation as quickly as possible for evidential purposes” [NP079]. A general duties officer identified the importance of timely attendance in situations where a risk of harm exists: “They're a risk-based assessment on the possible seriousness of the outcome – of failure to attend in a timely fashion to the incident.” [QG017]. A senior sergeant in charge of a police station district stated “I think probably one of the biggest things is the timeliness – being able to attend in a timely manner, because it's my view that that is of primary concern to members of the public.” [ND086]. These reasons, and the importance of a timely police response, are all identified in the literature reviewed in Chapter 3 (Bayley, 1994b; Eckblom & Heal, 1982; Herbert, 2006; Silberman, 1978). This finding shows that police personnel in both New Zealand and Queensland recognise the importance of timely response, and they understand why timely response is important.

**Apprehension of offenders**

Over 60% (n=74) of interviewees identified apprehension of offenders as important. Results for New Zealand Police and Queensland Police Service are similar. Nearly all general duties officers, but almost no performance managers or senior executives, identified apprehension of offenders as an important criterion of police effectiveness in responding to calls for service. This result demonstrates a clear difference between the perspectives of senior managers and general duties officers. This difference may reflect a change in the values or level of understanding that managers develop through their career.

Early literature assumed that rapid response to crimes by general duties officers influences the likelihood of apprehending offenders (Farmer, 1984; Larson, 1972a;
Manning, 1977), but more recent research shows that rapid response to most calls does not increase apprehension rates or reduce crime (Telep & Weisburd, 2012). The interviews of both senior executives and general duties officers identified rapid response as important. However, general duties officers identified apprehension of offenders more strongly than did senior executives, and senior executives identified demonstrating respect or empathy for the caller more strongly than did general duties officers. For example, when one senior executive was asked what they think matters in police responding effectively to calls for service (see Question 11 in Table 4.4 and Appendix B), this interviewee responded:

…it boils down to things like the empathy that we have to those that call us ….. the timeliness of response, whether or not there’s follow-through, whether or not we are actually presenting with people who are knowledgeable and know what they are doing [NS071].

**Prevention of harm or restoration of order**

Most interviewees (n=70) identified preventing harm or restoring order as important. For example, a general duties officer described the outcome of police responding to a particular call for service as being successful because “the offender/suspect was removed from the environment of which was causing quite a lot of distress at home” [NG066]. Junior personnel more readily identified this dimension as being important than did senior personnel. Over 70% (34 of 48) of interviewees more junior than the rank of sergeant, but less than 30% (7 of 24) of interviewees with a seniority of at least superintendent, identified preventing harm or restoring order as an important dimension of police effectiveness in responding to calls for service.
Attendance

Almost half of the interviewees (n=57) identified making the correct decision on whether to dispatch a unit to attend the job as important. Some difference existed between the two police agencies: 55% of interviewees from Queensland Police Service, but only 40% of interviewees from New Zealand Police identified the decision to dispatch a unit to attend a job as important. However, a more marked difference exists between different role types: three-quarters of communications room operators, but less than one third of general duties officers and performance managers, identified the decision to dispatch a unit to attend a job as important. This difference is unsurprising because the decision to dispatch a unit to attend a job is not made by performance managers and general duties officers, but is central to the role of communications room operators. For example, when asked if a particular job was successful one communications room operator responded that the job was successful because “I haven’t wasted sending a police resource” [QC004].

Literature reviewed in Chapter 3 identified that one challenge police face in responding to calls for service is coping with demand (Sparrow et al., 1990). Police cannot attend all calls for service, and must select the most important jobs for units to attend (Worden, 1993). The finding that many police personnel—particularly communications room operators—identify the decision to dispatch a unit to attend a job as important shows that police personnel who are involved in making these decisions understand the importance of such decisions in ensuring police are effective.

Satisfaction of the caller

Over 40% (n=51) of interviewees identified satisfying the caller as important. This criterion was expressed most strongly by PCC managers. Whereas some interviewees
expressed the importance of satisfying the caller as an outcome in itself, others saw it as important to avoid criticism. For example, one PCC manager said that it is important for police to do a thorough job because “then you can’t really criticise that” [NP102]. Prior research identifies the importance of first responders doing a thorough job to convict offenders (Forst, 1982), which is a dimension of police effectiveness from a technical-rational perspective. However, this PCC manager [NP102] was motivated to avoid criticism and maintain legitimacy, which is an objective explained by institutional theory (E. Davis, 2002; Herbert, 2006).

The five topics described in this section (timeliness of response, apprehension of offenders, prevention of harm or restoration of order, attendance, and satisfaction of the caller) are all identified in the previous literature reviewed in Chapter 3. However, analysis of the interviews revealed that senior managers have different perspectives from lower-ranked personnel on some dimensions of police effectiveness. Lower ranks more frequently identified technical-rational objectives, such as apprehending offenders and preventing harm, than did senior managers. However, senior managers more frequently identified objectives that support police organisational legitimacy, such as demonstrating respect for the caller. As discussed in Chapter 2, perceptions by an organisation’s employees of misalignment or confusion between technical-rational goals and institutional goals can lead to loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977). If junior personnel, who perceive of technical-rational goals as being important, are aware that senior personnel are more concerned with institutional goals, a condition that can cause loose coupling exists. Evidence of the existence of loose coupling in New Zealand Police and Queensland Police service will be explored in the next section.
5.4.2 Evidence of loose coupling in New Zealand Police and Queensland Police Service

Literature reviewed in Chapter 2 argues that loose coupling exists where agents are disconnected from organisational goals (Weick, 1990), and can be caused by administrative arrangements or by environmental factors such as the organisation’s social norms (Orton & Weick, 1990). Loose coupling manifests as a gap between formal structures and work practices. Formal rules are often broken, and decisions are often not implemented (P. Smith, 1993). In the previous section I found a difference between the perspectives of junior and senior police personnel. Interviewees who were junior in rank more frequently identified technical-rational objectives, such as apprehending offenders and preventing harm, than did senior managers. However, senior managers more frequently identified objectives that support police legitimacy, such as demonstrating respect for the caller. In the interviews I explored whether the different perspectives of junior and senior personnel reflect misalignment or confusion between technical-rational and institutional goals. In analysing the transcripts I addressed two sub-questions:

RQ3.1 Does evidence exist of loose coupling within police in which police employees perceive there to be misalignment or confusion between technical-rational goals and institutional goals?

RQ3.2 What influence does loose coupling have on the effectiveness of performance measurement frameworks?

Previous studies identified the presence of loose coupling within police agencies as a threat to the integrity of the performance measurement frameworks and therefore to police effectiveness (Andersson & Tengblad, 2009; Brunetto & Farr-Wharton, 2003, 2005;
Lawton et al., 2000; Lipsky, 1980; McKeivitt & Lawton, 1996). However, evidence on the existence and nature of loose coupling within police agencies is limited. In analysing the transcripts I explored three dimensions of loose coupling that are identified in the literature on police. These include subversion, which involves acting in ways contrary to the direction given by senior management (Brunetto & Farr-Wharton, 2005); disengagement, which manifests as lack of alignment with, and buy-in to, the objectives and mechanisms set by senior management (Lawton et al., 2000); and finally managers tolerating or ignoring breaches of policy by their staff (Shearing, 1984).

**Subversion of performance measurement**

Only two of the 121 interviewees provided examples of possible subversion of performance measurement. One general duties officer stated that senior managers use performance reports to “make themselves look good and justify the numbers of police” [QD027]. This interviewee stated: “I know at one stage here we had thirteen murders in the district, and they were called unlawful killings and changed. Within the statistical review we had one murder in the district, and I’m going: ‘Well, I attended four’.” The second example was provided by a performance manager who described an occasion when “a directive was issued by the District Commander about how they could improve their recorded performance as opposed to their actual performance” [NM062]. This interviewee described an example of police closing jobs in the dispatch management system (CARD) before attending the scene. This practice would have the effect of causing the response time performance measure to indicate that response time performance is better than it really is.
These examples accord with examples in previous research. Lawton et al. (2000) describe a British survey of a police management team that found “Anything that identified poor performance or inactivity that warranted investigation was excluded” (p. 17) from performance measurement. In a study of 178 Australian police officers, Brunetto and Farr-Wharton (2003) found police middle managers held negative attitudes following implementation of NPM reforms, and that “If first-level managers perceive a lack of synergy between a written policy and the supporting implementation variables (funding), then it is likely that, to the extent that they have power, they will use it to maintain the status quo” (p. 237).

Although the interviewees identified two examples of police possibly subverting performance measurement frameworks to make performance appear to be better than it really is, these are just two examples from 121 interviews. This is not evidence of a widespread practice of police personnel subverting performance measurement frameworks. However, these two examples both suggest that police want their measurement of performance to “make themselves look good” [QG017]. Other interviewees provided a similar perspective, which was discussed in Section 5.3 in the context of a theme of the influence of performance measurement frameworks.

**Disengagement and commitment of personnel**

Interviewees revealed wide variations in levels of engagement with the organisation’s objectives and processes. Most interviewees in both New Zealand and Queensland gave no evidence of disengagement. The least engaged staff often resent policies, systems, and management in general. However, these interviewees still appeared to demonstrate commitment to doing their job well, either in their own eyes or in the eyes of their...
immediate supervisor. For example, one senior executive described an experience that occurred while working in a previous role in a district. At that time, headquarters frequently changed the template for collecting performance data. This constant change caused staff to feel frustrated and disillusioned because the affected staff believed that constantly changing the template caused performance information to become meaningless. However, the interviewee observed about their staff that “they still earnestly did the best they can to provide the best responses” [NS060].

One communications room operator [QC013] who described himself as “a really disgruntled policeman” expressed disappointment about his police agency. This interviewee gave the reason for being disgruntled as: “I don’t believe that [citizens] are receiving a fair shake from the … government because I’m greatly disappointed with the management of resources within the service to try to answer calls for service.” The interviewee stated that this problem had existed for many years. As a result, he had lost confidence that the situation would ever improve. This same interviewee was concerned that managers do not care about police delivering an effective service, but only about avoiding criticism. “The only time that there’s any concern shown is if there is a complaint” [QC013]. This is clear evidence of misalignment or confusion between technical rational and institutional goals, which is a condition that can cause loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977). This perspective was expressed repeatedly across different roles in both jurisdictions. For example a communications room operator in a different police agency from interviewee QC013 stated: “We’ve become very much an arse covering workplace. In other words, we’re not using common sense; we’re using the fear of what might happen if something goes wrong, as opposed to using common sense in dealing with things” [NC103].
Although these examples reveal staff dissatisfaction with management, this dissatisfaction is not the result of a blue collar ethos concerned with conditions of service as described by Bayley (2008, p. 14). Instead, interviewees were frustrated with management decisions that they see as reducing police effectiveness. In particular, interviewees complained that police managers are more interested in avoiding complaints and managing their own career progression than in making lasting improvements. For example, when asked what they thought mattered most in police responding effectively to calls for service from the public, one first-line supervisor in a district said: “…leadership from above—which we have a definite lack of here—and enthusiasm and motivated supervisors which the section constables can see” [ND069]. Similarly, a communications room operator stated:

There’s no one here who wants to stay and hold the centre together, who actually wants to look at the issues and stay around long enough to implement changes and evaluate the effectiveness of them and then make further changes. People don’t stay around long enough to do that. [QC001]

Although interviewee QC001 believed their PCC does not evaluate its effectiveness, other interviewees did identify examples of police measuring their performance. However, some interviewees perceived performance measurement as ineffective because managers fail to act to improve performance when deficiencies are identified. In particular, these interviewees believed that management ignores poor performance and/or believed that if performance measurement was effective more resources would have been allocated. When asked for examples of consequences of performance being assessed as not good enough, one PCC manager stated: “To be honest, in the past I’ve seen staff get away with
some really terrible stuff … because we didn’t have any measures in place that sort of thing just got let go” [NP113]. This interviewee was hopeful that the recent introduction of performance measures would improve the situation. However, a manager in another PCC expressed disillusionment with performance measurement in police. This manager stated that performance measurement would “be helpful if I could get some more staff out of it, but no, it’s not” [QP012]. Similarly, a general duties officer stated: “As for general statistics, no it doesn’t help us in any way whatsoever. It doesn’t get us more resource, it doesn’t get us more manning, it doesn’t get us communications that work” [QG027]. A communications room operated stated: “It can only be helpful if someone does something when a deficiency is recognised, and it appears to me that the system is that they will look at things and recognise deficiencies but not act on them” [QC013].

Little evidence was found of senior managers communicating with staff to engender alignment with their vision and methods. Nevertheless, most interviewees expressed a desire to perform well and to provide a good service. Rather than looking to management to ensure their team work-groups are performing well, work groups self-regulate behaviours of their members through a culture of strong peer-pressure to work hard for the team. One first-line supervisor in a PCC stated: “I can’t stand people that don’t come to work to do their job properly; they frustrate me” [NP113]. Similarly, one first-line supervisor in a district stated: “Everybody knows who’s working; everybody knows who’s not working; everybody knows who’s lazy; everybody knows who’s out there doing what they need to do” [QD038].
Some interviewees—particularly in Queensland—expressed frustration because staff in districts often fail to log onto the dispatch system (CAD/CARD), or refuse to be dispatched to attend jobs. For example, in one interview a PCC manager expressed frustration because units that were supposed to be available for dispatch were unavailable: “We were chasing cars because we had a few jobs we needed to get responses to and we were ringing people that were supposed to be available and they were not, either not on, doing something else; there was no note on the sheet” [QP029]. Another PCC manager described how personnel in districts ignore a policy requiring all units on duty to log onto the CAD system and be available for dispatch. This manager stated that “the reality is that although policy exists, the practices on the ground haven’t changed dramatically” [QP044]. This interviewee described the police service as containing subcultures that have local values and a local work ethic that influences how personnel choose to work. Similarly, a district first-line supervisor said: “I find it frustrating that certain sections, like Traffic Branch who may be out there, just don’t respond to other jobs, or there are other sections that could attend those jobs, but because it’s not in their charter won’t attend” [QD052]. A communications room operator described the same problem: “It’s hard sometimes to get them to go; like you put the call up, but a lot of them—special units among them—it’s hard to get some people to go to jobs [sic]. It’s mainly the general duties that are doing all the hard yards” [NC074].

Consistent with the findings of Brunetto and Farr-Wharton (2005) and Shearing (1984), some managers I interviewed identified the existence of an organisational culture in which managers ignore breaches by first-line supervisors of policies. For example, one senior sergeant in a district stated: “Pretty much if the job gets finalised and there’s no
adverse comments back from it then we leave it and we don’t question to any degree, or any formalised structure, individual officers and what they’ve decided to do with jobs” [QD037]. Similarly, one PCC manager stated:

Most of the bosses have come up through that culture. Although the document and direction comes from the senior executive and they adhere to it, the implementation of that policy always rests with the supervisors and junior management, and they go back into the traditional practices, and the senior management will not take umbrage on that unless an issue actually occurs again. [QP044]

**Summary of findings about loose coupling**

In this section I addressed research question 3 and two sub-questions:

**RQ3 What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?**

**RQ3.1 Does evidence exist of loose coupling within police in which police employees perceive there to be misalignment or confusion between technical-rational goals and institutional goals?**

**RQ3.2 What influence does loose coupling have on the effectiveness of performance measurement frameworks?**

Interviews with personnel involved in delivering a service of responding to calls for service provide limited support for findings of prior research (Andersson & Tengblad,
2009; Brunetto & Farr-Wharton, 2003, 2005; Lawton et al., 2000; Lipsky, 1980; McKeVitt & Lawton, 1996) on the nature of loose coupling in police. Only two interviews out of 121 provided evidence of subversion of performance measurement. Most interviewees gave no evidence that they were disengaged with the vision, objectives or processes of the organisation. However, this section presented numerous examples of interviewees expressing frustration with management decisions that they saw as reducing police effectiveness. Some of these interviewees expressed disillusionment with performance measurement because of its failure to deliver improvements. The interviews provide some evidence of police personnel ignoring policies set by senior management, particularly in relation to personnel in districts failing to make themselves available for dispatch to attend calls for service. Some evidence was also provided of a culture in which managers tolerate such policy breaches. Rather than looking to management to ensure their work groups are performing well, work groups appear to self-regulate the behaviour of their members through a culture of strong peer pressure to work hard for the team. So, although this culture drives commitment by individuals to perform, these findings suggest that formal performance measurement frameworks in police may have little influence on the effectiveness of police.

As discussed earlier in the chapter, although both agencies routinely report quantitative performance information, I found few examples of such information being used in a consistent way across an agency to influence decisions. The interviews provide some evidence of PCCs monitoring demand volumes, call answer times, and response times. However, in districts, examples of managers using quantitative performance information are isolated and rare. If police decisions are rarely influenced by quantitative performance information, it follows that routine performance measurement will have negligible
influence on police effectiveness. To determine whether performance measurement does influence police effectiveness in responding to calls for service, I undertook a quantitative study described in the next chapter. In this study—Study 2—I examine the relationship between indicators of police effectiveness in responding to calls for service and performance measures monitored by police. Before forming specific testable hypotheses about this relationship I will review findings from Study 1 against prior research.

5.5 Moving Beyond Prior Research

Literature reviewed in Chapter 3 identifies numerous dimensions of police effectiveness in responding to calls for service, but provides little evidence on whether police personnel understand effectiveness in terms of these dimensions. Interviewees in Study 1 perceive of police effectiveness in responding to calls for service in similar terms to dimensions identified in the literature. In particular, interviewees identified timeliness of response, apprehension of offenders, prevention of harm, attendance, and satisfaction of the caller as dimensions of police effectiveness (see Section 5.4.1).

Interviewees who are junior in rank more frequently identified technical-rational objectives, such as apprehending offenders and preventing harm, than did senior managers. However, senior managers more frequently identified objectives that support organisational legitimacy, such as demonstrating respect for the caller (see Section 5.4.1). Previous literature argues that perceptions by an organisation’s employees of such misalignment or confusion between technical-rational goals and institutional goals can lead to loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977), and “successful implementation of policies requires that senior and lower managers must be in congruence in relation to the stated goals and objectives of a new policy” (Brunetto &

Literature on loose coupling in police identifies impacts of loose coupling that include subversion of performance measurement (Brunetto & Farr-Wharton, 2005), disengagement of employees from the organisation’s objectives and methods (Lawton et al., 2000), and tolerance by managers of breaches of policy by their staff (Shearing, 1984). Only two of the 121 interviews in Study 1 revealed examples of police apparently subverting performance measurement to make performance appear to be better than it really is. Most interviewees gave no evidence that they were disengaged with the vision, objectives or processes of the organisation. However, some expressed frustration with management decisions that they saw as reducing police effectiveness. Some interviewees expressed disillusionment with performance measurement because of its failure to deliver improvements (see Section 5.4.2).

The narratives of interviewees provide some evidence of police personnel ignoring policies set by senior management, particularly in relation to personnel in districts failing to make themselves available for dispatch to attend calls for service. Some evidence was also found of a culture in which managers tolerate such policy breaches. Rather than looking to management to ensure their work groups are performing well, work groups appear to self-regulate the behaviour of their members through a culture of strong peer pressure to work hard for the team (see Section 5.4.2). So, although this culture drives commitment by individuals to perform, if police personnel look to peers rather than
support objectives and performance criteria specified by management, performance measurement frameworks established by management may have little influence on the effectiveness of police (Lawton et al., 2000).

Literature reviewed in Chapter 2 suggests that police closely monitor measures related to crime but ignore the many other things that police do (Fleming, 2009; Kelling, 1992; Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; M. H. Moore & Braga, 2003; Skogan & Frydl, 2004). These scholars argue that the absence of broader performance measures prevents police from being as effective as they could be. The narratives of police personnel interviewed in Study 1 identify performance measures related to crime, such as offence volumes and crime resolution rates. However, these narratives also identify measures that relate to both crime and non-crime response activities of police, such as response timeliness, complaints, customer satisfaction, noise in the media or parliament, volumes of calls for service, and volumes of repeat calls for service to the same address (see Section 5.3.4).

Some scholars suggest police managers may lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008). Findings of Study 1 support this suggestion: Although performance measures are routinely reported throughout both New Zealand Police and Queensland Police Service few interviewees appear to use performance information effectively. Instead, the dominant approach identified by interviewees for managing performance involves checking the work of subordinates (see Section 5.2).

These findings, taken together, suggest that performance measurement in police may have little influence on the effectiveness of police in responding to calls for service. To test
whether this is the case I undertook a second study that examines the relationship between measures monitored by police and police effectiveness. Tested in Study 2 are the hypotheses that arise from the findings of Study 1.

5.6 Generating Hypotheses for Study 2

So far in my thesis I have taken three steps to gather information about the nature of the relationship between performance measurement in police and police effectiveness in responding to calls for service. First, I reviewed literature on performance measurement in the public sector in general and in police in particular. Second, I reviewed literature on police effectiveness in responding to calls for service. Third, I explored narratives of personnel in New Zealand Police and Queensland Police service who are involved in delivering or managing the service of responding to calls for service from citizens.

Together these sources aid in the development of hypotheses about the relationship between performance measurement in police and the effectiveness of police in responding to calls for service. These hypotheses arise from propositions that emerge from prior research and findings of Study 1, and are based on the conceptual model presented in Figure 5.1. Central to this model is the idea—illustrated by arrow (a) in the diagram—that performance measurement can directly assist the organisation to be effective in technical-rational terms. For example, managers monitor performance measures related to service delivery to inform decisions that support the effective and efficient delivery of service (Hoque et al., 2004; R. W. Scott, 1987). However, as illustrated by arrow (b), rather than use performance measurement to influence decisions, managers may adopt performance measurement to enhance their organisational legitimacy by appearing progressive and successful (Willis et al., 2007), and by demonstrating structural
conformity (isomorphism) with other organisations (DiMaggio & Powell, 1983; Willis et al., 2007). Performance measurement acts as a buffer to allow internal actions to vary in response to practical considerations (Meyer & Rowan, 1977; Willis et al., 2007), shown as arrow (c).

**Figure 5.1**
The Influence of Performance Measurement on Police Effectiveness

Where performance measurement is not used to influence decisions in technical-rational terms, employees perceive misalignment or confusion between technical-rational goals and the institutional goals of managers, which creates loose coupling (Lawton et al., 2000; Meyer & Rowan, 1977), shown as arrow (d). This loose coupling includes disenchantment with performance measurement (see Section 5.4.2), which further lessens—shown as arrow (e)—the ability of performance measurement to directly influence technical-rational outcomes (Collier, 2001; P. Smith, 1993). Even though senior managers impose performance measurement mechanisms on the organisation,
performance measures that police monitor may have little influence on the decisions and behaviours of middle managers and front-line staff, who have become decoupled from senior management (Brunetto & Farr-Wharton, 2003, 2005; Meyer & Rowan, 1977; Taylor, 2009).

Loose coupling is not the only mechanism affecting the influence of performance measurement on police effectiveness. Shown as arrow (f), police managers may lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008). Study 1 found that, although police measure their performance, police managers make only limited use of quantitative management information. So, although Study 1 found that police monitor both measures of service delivery performance and indicators of police legitimacy, loose coupling and a lack of skill in using performance information may prevent this monitoring from influencing police effectiveness. To test these ideas, I generate four testable hypotheses.

In Study 1 I identified various measures that New Zealand Police and Queensland Police Service include in OPRs and routine performance reporting. In this section I outline hypotheses that test the relationship between these measures and outcomes of police effectiveness in both technical-rational and institutional terms. These hypotheses are:

Hypothesis 1. The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service

Hypothesis 2. Police do not respond to performance measures in a way that influences their organisational legitimacy
Hypothesis 3. Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.

Hypothesis 4. Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

My first three hypotheses concern the influence of performance measurement on police effectiveness in responding to calls for service. These three hypotheses together help to test an overall hypothesis that police managers do not respond to quantitative performance information in a way that influences police effectiveness. In Study 1 I also found that police personnel who are disillusioned with management and/or performance measurement still maintain a commitment to working hard and doing a good job in their own eyes or in the eyes of their work-group. My fourth hypothesis tests whether police continue to be effective when employees are demotivated. These four hypotheses will be described.

5.6.1 Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.

Prior research does not provide a clear picture of whether performance measurement helps or hinders police effectiveness. Loose coupling theory suggests performance measurement may be counter-productive, driving staff disengagement and even subversion of policy or the performance that is reported (Brunetto & Farr-Wharton, 2003; Lawton et al., 2000). Most interviewees in Study 1 gave no evidence they are subversive or disengaged with the organisation. However, in Study 1 I found some evidence of loose
coupling in which some managers tolerate policy breaches by their employees, and some employees feel disillusioned with management. One reason for this disillusionment is that these police personnel perceive performance measurement in police to be ineffective (see Section 5.4.2).

Theory behind NPM suggests performance measurement can improve police performance in delivering service to citizens (O’Flynn, 2007). In Study 1 I identified numerous performance measures monitored by police. However, it found few examples of such information being used in a consistent way across an agency to influence decisions. Some scholars suggest police managers may lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008). In the narratives from Study 1, examples of managers using quantitative performance information in police districts are isolated and rare. If police decisions are rarely influenced by quantitative performance information, it follows that routine performance measurement has negligible influence on measures of police effectiveness; variations in performance measures monitored by police will not influence police effectiveness in responding to calls for service.

Hypothesis 1 considers the influence of performance measurement overall. However, two distinct purposes for performance measurement are identified in both the literature and in Study 1. First, to deliver technical-rational outcomes: Measuring performance enables members of an organisation to make sense of what is happening in complex organisational contexts (Gill, 2011), and informs management decisions that support effective and efficient delivery of services (Hoque et al., 2004; R. W. Scott, 1987). Second, to deliver institutional outcomes: Performance measurement frameworks act as legitimising
structures (Taylor, 2009), and provide a buffer that enables internal actions and decisions to vary in response to practical considerations (Meyer & Rowan, 1977, p. 357). I therefore hypothesise:

5.6.2 **Hypothesis 2: Police do not respond to performance measures in a way that influences their organisational legitimacy**

Prior research in the U.S.A. found evidence of police departments adopting a performance measurement framework (Compstat) to appear progressive and successful. Accordingly, Compstat was implemented in a way that would confer legitimacy and minimise disruption to existing routines (Willis et al., 2007). That study further concluded that Compstat “was less successful when trying to change those structures and routines widely accepted as being ‘appropriate’” (p. 147). Consistent with the findings of Willis et al. (2007), interviewees in Study 1 who expressed negative perceptions of performance measurement in police were critical of the inaccuracy of performance measures or the ineffectiveness of performance measurement to deliver technical-rational outcomes such as sufficient resources or improvements in service. However, like Willis et al. (2007), the results of Study 1 provided evidence of managers using performance measurement frameworks to enhance organisational legitimacy, either as a reaction to public criticism, or by monitoring indicators of organisational legitimacy, rather than of technical-rational measures. It follows that police monitoring of performance measures may influence measures of organisational legitimacy. However, the results of Study 1 and the literature (Andersson & Tengblad, 2009; Coleman, 2008) also suggest police managers may lack the skills to use performance information effectively. It may be that police operate
performance measurement frameworks to appear legitimate, rather than to improve their effectiveness (Taylor, 2009).

5.6.3 Hypothesis 3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service

In Study 1 I found evidence of police managers monitoring complaints and noise in the media and parliament to determine whether police are performing effectively. If these police managers act in response to variations in levels of complaints and noise this may influence police effectiveness in technical-rational terms. However, findings of Study 1 also show that police managers make limited use of quantitative performance information of any kind. Instead, they tend to rely more strongly on checking individual instances of work and on reviewing individual instances of police being criticised, either through complaints or noise in parliament or the media. I therefore hypothesise that variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.

If Study 2 finds that Hypotheses 2 and 3 are false this will not only support the proposition that police measure performance to enhance their legitimacy, but will also challenge the proposition in the literature that police managers lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008). If, however, findings of Study 2 support Hypotheses 1, 2, and 3, this will reinforce the propositions in the literature that police managers either lack the skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008), or measure performance to
appear legitimate, rather than to improve service delivery (Taylor, 2009; Willis et al., 2007).

5.6.4 Hypothesis 4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take

In Study 1 I found that police personnel who are disillusioned with management and/or performance measurement still maintain a commitment to working hard and doing a good job in their own eyes or in the eyes of their work-group. Although the results of Study 1 provide insight on the perspectives of police personnel, Study 1 did not measure whether these perspectives translate to behaviour that maintains police effectiveness in responding to calls for service even when employees are disillusioned with management. Organisations typically use surveys to monitor levels of staff engagement (Attridge, 2009). However, such surveys are costly and typically occur infrequently. Disengaged staff are likely to have lower morale than engaged staff. A proxy indicator of morale, such as the amount of sick leave taken (Bernberg, 1952; Parker, 1995), is readily available from police administrative data. I therefore hypothesise that in Study 2, police effectiveness in responding to calls for service will be insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

5.7 Summary

Drawing on my review of prior literature and the narratives of police personnel involved in responding to calls for service from citizens, I generated four hypotheses about the relationships between police use of performance information and effectiveness. These hypotheses arise from propositions that emerge from prior research and the findings of
Study 1. These propositions are that police do not act in response to performance measures that they monitor in a way that influences police effectiveness, either in technical-rational or institutional terms, and that police personnel continue to do their best to do a good job even when disillusioned with management and the performance frameworks imposed by management.

Specifically, I argue that police managers either lack the incentives or skills to use performance information effectively, and that the resultant loose coupling preserves the effectiveness of the organisation due to an organisational culture characterised by strong peer-pressure to work hard and do a good job for the team. This argument arises from the sources of information discussed so far in this thesis: literature on performance measurement in police, literature on police response to calls for service, and narratives of personnel involved in delivering or managing the service of responding to calls for service from citizens. To substantiate my argument, I undertook a second study that tests the hypotheses outlined in the previous section. If findings of Study 2 show that police do respond to performance measures in a way that improves measures of police effectiveness, my argument will be weakened. If, however, the findings from Study 2 find no evidence that police respond to performance measures in a way that improves measures of police effectiveness, my argument will be strengthened.
Chapter 6: Study 2 Method

6.1 Introduction

Moynihan and Pandey (2010) argue “that understanding public employee use of performance information is perhaps the most pressing challenge for scholarship on performance management” (p. 849) because of the effort governments have given to creating performance data to improve governance, and because there is much we do not know about the factors associated with the use of that information. Yet this relationship is not well understood in the literature. In Study 1 I explored how police personnel use performance information. I also explored the perceptions police personnel have about the influence of performance measurement on police effectiveness in responding to calls for service. Based on the results of Study 1 I identified various distinct dimensions of police effectiveness. Some of these dimensions reflect effectiveness in technical-rational terms, such as service delivery; others reflect effectiveness in institutional terms, such as organisational legitimacy. The results of Study 1 provide examples of police managers using performance information to improve service delivery. The study also found evidence of police seeking to enhance their organisational legitimacy through measuring and reporting quantitative performance information.

However, in Study 1 I found only limited evidence of managers being influenced by quantitative performance information. Instead, the predominant mechanism police managers use to monitor performance is qualitative review of individual instances. Therefore, the performance measures that police monitor may have little influence on police effectiveness. The results of Study 1 also identified the presence of loose coupling. Some police personnel are disillusioned with senior management and performance
measurement. However, these personnel appear to maintain a commitment to working hard and doing a good job. These findings suggested four hypotheses about the nature of the relationship between police use of performance information, morale, and effectiveness. These hypotheses are:

H1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.

H2: Police do not respond to performance measures in a way that influences their organisational legitimacy.

H3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.

H4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

To test these hypotheses, in Study 2 I acquire and analyse time series data from both New Zealand Police and Queensland Police Service on a range of variables identified in Study 1. I begin this chapter by outlining the rationale for the method employed in Study 2. I then describe the steps in this method: select variables, acquire data, pre-whiten time-series, cross-correlate time series, and perform Granger-causality testing.
6.2 Selection of Method

Performance measurement processes involve measuring numerous distinct indicators, known as performance measures, from a recent historical period to inform decisions affecting future periods. If performance measurement processes influence future periods then one would expect cross-correlation of time series to reveal a relationship between relevant variables in which the independent variable precedes the dependent variable in time. It follows that a bivariate, or possibly multivariate, time series analysis approach would be a suitable methodology to test such relationships. A multivariate analysis would be appropriate if police look at multiple distinct performance measures in combination to influence decisions. Interviewees in Study 1 provided no evidence of using performance measures in combination. However, interviewees did identify a number of distinct measures as indicators of police effectiveness (outcomes) or measures that police monitor to influence their decisions (inputs). I therefore elected to use a bivariate time series analysis approach.

Time series analysis can be a strong quasi-experimental design (T. D. Cook & Campbell, 1979; Schwert, 1979). However, to avoid producing spurious results with time series analysis, technical challenges must be overcome (Ashley, Granger, & Schmalensee, 1980; Erol & Yu, 1987; Schwert, 1979). These challenges will be described.

Traditionally the most common method for describing the relationship between two time series’ has been through a lagged cross-correlation function (Haugh & Box, 1977; Schwert, 1979). The main challenge with this approach is controlling non-stationarity and autocorrelation, which can affect values of the cross-correlation function. Non-stationarity and autocorrelation can erroneously give significance to apparent patterns in
the cross-correlation function (Ashley et al., 1980; Hanssens, Parsons, & Schultz, 2003; Pfaff, 2008; Schwert, 1979). This problem can be addressed by removing autocorrelation, trend, and seasonality, thereby transforming the time series to white noise (Loftin & McDowall, 1982; McCleary & Hay, 1980; Wei, 2006). Time series analysts have used various approaches to transform time series to white noise, including spectral analysis, and fitting econometric models or ARIMA models to the time series to filter out deterministic patterns. Helmer and Johansson (1977) explain these approaches, and find ARIMA models (Box & Jenkins, 1976) to be superior to approaches based on econometric models or spectral analysis.

Loftin and McDowall (1982) provide an example of using a bivariate ARIMA model approach to determine the relationship between numbers of police officers and crime. In this approach they pre-whitened each time series by fitting an ARIMA model to each time series—an approach known as double pre-whitening (Haugh & Box, 1977). Loftin and McDowall (1982) then computed the cross-correlation function of the resultant residuals. This revealed the relationship between these two variables. Finally, by performing Granger-causality tests, they tested whether or not this relationship was causal. Loftin and McDowall (1982) argue that it is not strictly true that correlation between two variables does not imply causation: “If non-linearities and third variable causation can be ruled out or controlled, correlation can indeed be interpreted in causal terms” (p. 395).

Although it can be difficult to rule out the possible influence of an unknown or hidden variable, it is often possible to determine whether one variable can help predict another. Granger (1969) developed an approach to address the question of whether $x$ causes $y$: by quantifying how much of the value of $y$ at a given point on time can be explained by past
values of $y$, then seeing whether adding lagged values of $x$ can improve the explanation. This has become a standard approach for causal analysis using time series (Horvath, 2003; Wei, 2006). $Y$ is said to be Granger-caused by $x$ if $x$ helps in the prediction of $y$, or equivalently if the coefficients of the lagged $x$’s are statistically significant (Schwert, 1979). Granger-causality is not a complete proof of causality, but it is a strong indicator of possible causality and “incremental predictability” (Schwert, 1979, p. 82) between time series.

Methods involving pre-whitening, computing cross-correlation functions of residuals, and performing Granger-causality tests are frequently referred to as transfer function analysis. This has become the standard methodological approach for analysing the relationship between two time series’ (Schwert, 1979; Wei, 2006). I therefore elected to use transfer function analysis in this study.

To create a robust ARIMA model with sufficient information to detect a significant relationship, if one exists, a time series should contain at least 50 observations (McCleary & Hay, 1980). Powerful algorithms such as TRAMO (Gomez & Maravall, 1997) have been created to help determine a suitable structure for an ARIMA model that will control for seasonality of monthly data. I therefore elected to use time series involving data points at monthly intervals for a minimum five year period between July 2004 and June 2009. This approach ensured (a) each time series contained at least 60 data points, and (b) that TRAMO could be used to help determine suitable ARIMA models.

Having selected a method, I followed the example of Loftin and McDowall (1982) by applying the following five steps:
1. Identify, from the results of Study 1, variables representing relevant dimensions of effectiveness and measures that police monitor (performance measures).

2. Obtain time series data for these variables.

3. Transform (pre-whiten) all time series’ to make them suitable for cross-correlation analysis.

4. Calculate cross-correlation functions between relevant pairs of time series’, to determine whether a relationship exists between the variables and what time lag is evident in this relationship.

5. Perform Granger-causality tests on these bivariate pairs, to determine whether a given ‘x’ variable can help predict the relevant ‘y’ variable.

These steps are described in Sections 6.3 to 6.7 below.

6.3 Step 1: Select Variables

Numerous measures relevant to the hypotheses to be tested were identified in Study 1 and the literature. These include three measures that New Zealand Police and Queensland Police Service monitor in their routine monthly performance measurement processes (input variables): demand volumes (DV), offence volumes (OV), and emergency response time performance (ER); three measures of police effectiveness in technical-rational terms: general response time performance (GR), repeat calls for service (RC), and resolution rate (RR); and two measures of police effectiveness in institutional terms: complaint volumes (CV) and noise (NV). The amount of sick leave taken (SL) is a proxy
indicator for staff morale (Bernberg, 1952; Parker, 1995). Finally, although not a performance measure, interviewees in Study 1 identified staff headcount (SH) as providing police with the capacity to be effective in responding to calls for service, and as something these interviewees believe would be influenced by performance measurement processes if these processes were effective. These 10 variables will be described.

**Input variable: Demand volumes (DV)**

The literature identified that one challenge police face in responding to calls for service is coping with demand (Sparrow et al., 1990). Both New Zealand Police and Queensland Police Service operate a computer aided dispatch system (CARD/CAD) in which information collected during jobs of responding to calls for service is recorded (see Chapter 4). These systems can report the number of jobs recorded as a result of calls for service to police. Interviewees identified that both agencies monitor the volumes of these jobs as a measure of demand. Interviewees perceived that police monitor demand volumes to influence decisions about how many staff to assign to the part of the agency that must cope with that demand. The literature also identified coping with demand as a challenge for police (Sparrow et al., 1990). I therefore selected DV as an independent variable to use when testing hypotheses 1 and 2.

**Input variable: Offence volumes (OV)**

The literature identified crime statistics as performance measures that police monitor (Maguire, 2003). Interviewees from both New Zealand Police and Queensland Police Service routinely monitor volumes of recorded offences as part of their performance measurement frameworks. Because police monitor offence volumes to influence
decisions I selected OV as an independent variable to use when testing hypotheses 1 and 2. Police categorise offence volumes into different types of crime, each of which is reported separately in performance measurement frameworks. The type of crime most commonly referred to by interviewees is burglary which, under the Australian and New Zealand Standard Offence Classification (ANZSOC), is referred to as Unlawful Entry With Intent.

Many crimes go unreported to police. For example, victimisation surveys typically show that most victims of sexual offences or domestic assault do not report these offences (Morrison et al., 2010). Consequently, variations over time in volumes of these types of offences, or even of total recorded crime, may be confounded by changes in reporting rates. Serious property crime, such as burglary or theft of motor vehicles, is usually reported to police (Morrison et al., 2010). Victims are often motivated to call police, either because they hope police will be able to recover their property, or because insurance companies require a police report before they will accept a claim (Morrison et al., 2010). So, because burglary features strongly in police performance measurement frameworks and because most burglaries are reported to police, I included volumes of burglary offences, rather than total crime, in this study.

**Input variable: Emergency response-time performance (ER)**

Police in both New Zealand and Queensland prioritise jobs for dispatch. In both agencies the jobs police give high priority for rapid response are those that involve either immediate threat to life or property or offenders still being present at the scene. Such jobs may also pose greater risk to police organisational legitimacy if response is slow (Moskos, 2007; Sparrow et al., 1990). Failing to respond rapidly to an emergency situation may
expose police to public criticism. The performance measurement frameworks in both New Zealand Police and Queensland Police Service include measures of police performance in responding to emergency jobs. I therefore include emergency response-time performance (ER) as an independent variable in Study 2.

Technical-rational effectiveness indicator: General response-time performance (GR)

Citizens do not only expect police to respond in a timely way to emergencies; they like to believe that the police will arrive quickly if they are called (Eckblom & Heal, 1982). Citizens’ perceptions of the speed of police response to calls for service have been shown to be a strong predictor of satisfaction with police performance (Parks, 1984; T. Pate, 1976; Percy, 1980). Although not monitored as closely in performance measurement frameworks as ER, both New Zealand Police and Queensland Police Service prioritise other calls for service as requiring a timely response. Unlike emergency jobs, these jobs are not treated as an urgent response requiring flashing lights and sirens. They are the next priority tier, where the PCC assesses that it is important that police respond in a timely way when a unit becomes available from another job. Whereas emergency jobs make up a negligible proportion of police calls for service, this next tier of jobs requiring timely response represents the majority of calls for service to police. Response time to these jobs (GR) is therefore a better indicator of overall police response time performance than is emergency response-time performance (ER). I therefore include general response-time performance (GR) as an indicator of police effectiveness.
**Technical-rational effectiveness indicator: Repeat calls for service (RC)**

Whereas offence volumes (OV) can be influenced by factors unrelated to the police, the literature identified calls for service as opportunities to reduce repeat victimisation (Farrell & Pease, 1993; Pease, 1998; Townsley et al., 2000) and thereby reduce crime. The volume of repeat calls for service therefore represents a measure of police effectiveness. Many police personnel interviewed in Study 1 identified the prevention of Repeat Calls for service as a dimension of police effectiveness in responding to calls for service. Some interviewees described solving the problem so that police do not need to be called back in the future as a measure of successful service delivery by police. I therefore include RC in Study 2 as an indicator of police effectiveness in technical-rational terms.

**Technical-rational effectiveness indicator: Resolution rate (RR)**

Interviewees in Study 1 identified resolving crime as an important dimension of police effectiveness. However, the volume of resolved offences may not be a suitable variable to include in this study because it may be correlated with the volume of recorded offences simply because each offence provides opportunities to resolve. Resolution rate (the proportion of recorded offences that police resolve) is a more obvious measure of police performance in technical-rational terms, because it is an indicator of the success rate in resolving crime. Because burglary offences rather than total crime are included in this study I include burglary resolution rate (RR) rather than resolution rate for all crime types. By aligning the group of offences included in OV and RR I avoid diluting the strength of relationship between these two variables, which would occur if offences were included in one variable that that are excluded from the other.
In Study 1 the most commonly identified measure of police effectiveness or performance was complaints. Police personnel are very aware of the potential for complaints being made, and they strive to avoid complaints. A complaint by a member of the public about police indicates a level of dissatisfaction. Complaint volumes (CV) therefore represents a measure of police organisational legitimacy in institutional terms.

**Institutional effectiveness indicator: Complaint volumes (CV)**

Many interviewees, particularly in New Zealand Police, monitor issues appearing in the media or parliament that reflect negatively on police (see Section 5.3.3). These interviewees consider the amount of ‘noise’ in the media to be a useful and/or important indicator of police performance. Noise in the media and noise in parliament were often mentioned together as being similar indicators, because issues in the media often generate debates in parliament. Like complaints, noise in the media or parliament is seen as a negative indicator of, or even a threat to, police organisational legitimacy. Police therefore strive to avoid noise, which represents a measure of police effectiveness in institutional terms.

**Institutional effectiveness indicator: Noise (NV)**

To test Hypothesis 4, I wish to test whether police maintain their effectiveness when police personnel are disillusioned with management and have low morale. As discussed in Section 5.6, employee morale is difficult to measure directly. The amount of time employees take off work on sick leave has for many years been one of the most common indicators of employee morale. It is frequently used in research and in management information as an indicator of employee morale (Bernberg, 1952; Parker, 1995). The
human resources systems of both New Zealand Police and Queensland Police Service are able to report how much sick leave personnel take during a given period. In Study 2 I therefore use sick leave data from these human resources systems to provide a proxy measure for employee morale.

**Staff headcount (SH)**

In Study 1, the predominant reason interviewees gave for police measuring demand volumes is to plan resource allocation; specifically staff headcount. Although staff headcount is not a measure of police effectiveness, interviewees perceived it as representing capacity to deliver the service of responding to calls for service. Therefore testing the relationship between input variables DV, OV, and ER (as independent variables) and SH (as a dependent variable) will test whether police respond to demand volumes in the way suggested by interviewees. This will help to test Hypothesis 1.

Some interviewees believed that having more staff will help police to cope with demand volumes and thereby be more effective. I therefore also test the relationship between SH (as an independent variable) and the measures of police effectiveness in responding to calls for service (GR, RC, and RR). I perform this test to show whether variations in staff headcount affect police effectiveness in responding to calls for service in technical-rational terms (Hypothesis 1).

**6.4 Step 2: Acquire Data**

I am fortunate in that both New Zealand Police and Queensland Police Service were willing to supply administrative data from relevant IT systems. Four systems from each agency contain relevant data, as follows:
• CSS (Queensland) / IAPRO (New Zealand) – Used to record and manage complaints made against police. These systems contain data on complaint volumes (CV).
• CAD (Queensland) / CARD (New Zealand) – Used by PCCs to manage the initial response to calls for service. They contain data on demand volumes (DV), emergency response time performance (ER), general response time performance (GR), and repeat calls for service (RC).
• QPRIME (Queensland) / NIA (New Zealand) – Used to record crime, investigations, and resolutions. They contain data on offence volumes (OV) and resolution rate (RR).
• AURION (Queensland) / PEOPLESOFT (New Zealand) – Human resources systems. These systems contain data on staff headcount (SH) and sick leave (SL).

Neither Queensland Police Service nor New Zealand Police operate an IT system containing suitable data on noise (NV). It was therefore necessary to construct a time series from publicly available data. A description of the data on noise (NV) and the data sourced from each police agency follows:

CSS / IAPRO

CSS contains a record for each complaint made against a Queensland Police Service employee. One complaint may contain more than one specific allegation. CSS categorises each complaint according to the location of the employee, the date the complaint was received, the date of the event to which the complaint relates, and the rank and role of the police employee. CSS further categorises each allegation according to the type of behaviour that has been alleged as being inappropriate. However, CSS does not identify
the context from which the complaint arises. So, it cannot isolate complaints arising from police response to calls for service.

IAPRO performs a similar function in New Zealand Police to the function performed by CSS in Queensland Police Service. However, some differences exist. IAPRO does not distinguish between complaints and allegations; neither does it contain as many allegation categories as CSS. However, unlike CSS, IAPRO does record the type of service that the police officer whom the complaint relates to was delivering. For consistency, I elected to count all complaints against police in a given district. However, due to differences in information system life-cycles, data were available for slightly different periods, as follows:

- From CSS, the number of complaints received during each month from April 2004 to December 2009 inclusive. (69 data points)
- From IAPRO, the number of complaints received between July 2003 and June 2009 inclusive. (72 data points)

**CAD / CARD**

CAD and CARD contain records of all calls for service jobs recorded by PCCs. These systems include information about these jobs, including the type of job, its priority for attendance, and various time-stamps, such as the time the job is first recorded, and the time a unit arrives at the scene. For both agencies, I was able to obtain data for all months from July 2003 to June 2009 inclusive (72 data points), as follows:

- DV – Counts of all calls for service jobs recorded by PCCs.
• ER – The proportion of emergency calls for service that police succeed in attending within 10 minutes.

• GR – The proportion of non-emergency calls for service requiring timely response that police succeed in attending within one hour.

Queensland Police Service was unable to supply data on repeat calls for service, but New Zealand Police did supply numbers of repeat calls for service (RC), counted as follows:

Count all jobs recorded in CARD where police were called to an address in the specified month, and where police had been called to that same address on a previous occasion in any of the specified month or the two preceding months.

**QPRIME / NIA**

QPRIME and NIA contain records of all offences recorded by police. These records usually result from calls for service recorded in CAD / CARD, but may be categorised differently than in CAD / CARD. For example, police may determine after investigation that what actually happened differs from what was reported to the PCC. In such instances the offence code given to the job in QPRIME / NIA will be changed from the offence code that was initially given to the job in CAD / CARD.

QPRIME and NIA also identify which recorded offences resulted in alleged offenders being apprehended. When one or more alleged offenders are recorded against a recorded offence police refer to this as an offence clearance. This enables Resolution-rate (RR) to be calculated by dividing the number of offence clearances recorded during a given month by the number of offences recorded by police during that month.
For both agencies I obtained data for all months from July 2003 to June 2009 inclusive (72 data points), as follows:

- **OV** – Counts of all recorded offences of burglary
- **RR** – The number of burglary clearances recorded during a given month divided by the number of burglaries recorded by police during that month

**AURION / PEOPLESOFT**

AURION and PEOPLESOFT are the human resources ICT systems used by Queensland Police Service and New Zealand Police respectively. Both agencies provided the numbers of sick days taken by staff during any given month (SL). New Zealand Police also supplied monthly figures for staff headcount (SH).

Queensland Police Service provided sick leave (SL) data for January 2004 to December 2009 (72 data points). NZP provided sick leave (SL) data for July 2004 to April 2010 (70 data points). NZP also provided staff headcount (SH) data for July 2003 to April 2010 (82 data points).

**HANSARD**

Although interviewees in Study 1 frequently mentioned noise in the media or parliament as an indicator of police performance or effectiveness, neither agency has data from which a time series of noise (NV) could be formed. To form a meaningful time series it is necessary to have a consistent measurement frame over time. Noise in the media is too unstructured and diverse for a consistent measurement frame to be feasible. However, the state parliaments of both New Zealand and Queensland record parliamentary debates
verbatim in *Hansard* (Hansard, 2014). So, although tedious to code, this does provide a data source from which meaningful time series can be formed.

Because of the time and effort involved to manually code data from Hansard, I elected to create time series from one parliament only. Because New Zealand Police was able to supply more data than Queensland Police Service, a time series of New Zealand Hansard data would be able to be cross-correlated with more other variables than would a time series of Queensland Hansard data. I therefore elected to derive time series from New Zealand Hansard.

I employed a research assistant to search five years’ of parliamentary debates as recorded in Hansard. I asked the research assistant to identify each distinct instance of police being criticised or spoken about in unfavourable terms during parliamentary debates. After checking data produced by the research assistant, I formed a time series (NV) containing counts of all issues raised in parliament that were critical of police. These issues mainly relate to police integrity, performance, or resourcing. The resultant time series contains 60 monthly data points. These data points correspond to all months from July 2004 to June 2009 inclusive.

Because parliament operates at state level, and many debates about police referred in general to police rather than to a specific police district, it was not possible to form a separate time series for each district. Instead, I created one single time series for New Zealand debates, and used this in my analysis of each district.

### 6.5 Step 3: Pre-whiten Time Series

From the above data I created 49 time series for inclusion in this analysis, as follows:
• Nine time series (DV, OV, ER, RR, GR, RC, CV, SL, and SH) for each of the three New Zealand Police districts
• Seven time series (DV, OV, ER, RR, GR, CV, and SL) for each of the three Queensland Police Service districts
• One national level time series for New Zealand from Hansard (NV)

Having created these time series I performed Step 3 of the method. This involved transforming (pre-whitening) time series to make them suitable for cross-correlation analysis. I used EViews (version 7) to perform this pre-whitening. EViews is a statistical programme that is commonly used for econometric time series analysis. It is particularly suitable for use in this study because it contains many relevant features to the analysis required. These include TRAMO, Granger-causality testing, cross-correlation, a flexible ARIMA modelling capability, a flexible Vector Auto Regressive (VAR) modelling capability, outlier detection, extensive graphical functionality, and generation of test-statistics to assist in assessing how well a given model fits the data.

Some scholars have criticized the double pre-whitening method for rejecting the null hypothesis of independence too frequently (Erol & Yu, 1987; Schwert, 1979). Hanssens et al. (2003) and Schwert (1979) recommend carrying out tests involving double pre-whitening with care, and supplementing results with other analysis. Therefore, to reduce the potential for producing spurious results, rather than only fit an ARIMA model to each time series, I also fitted two Vector Auto Regressive (VAR) models to each time series. The particular statistical properties of a given time series affect how well an ARIMA model can fit the data. Different types of time series model may fit a given time series better than others. Although ARIMA models usually perform better than VAR models
(Helmer & Johansson, 1977), depending on the particular statistical properties of the data, VAR models sometimes perform better (Bagshaw, 1987).

The two VAR models I fitted to each time series differ in the complexity of the model through the number of lags they incorporate. For one model I applied Akaike’s Information Criterion (AIC), which den Haan & Levin (2000) showed optimises the quality of the model by balancing the trade-off between goodness of fit and complexity. In my second VAR model I fixed the number of lags at 12, this being the length of a full annual cycle. This approach produces an over-fitted model compared to the ARIMA model and the VAR-AIC model. An over-fitted model is less likely to produce Type 1 errors, but more likely to produce Type 2 errors. However, on average, resultant residuals resemble white noise more closely.

6.6 Step 4: Cross-correlate Time Series

Having, in Step 3, separately applied ARIMA and VAR models to pre-whiten each time series, in Step 4 I calculated cross-correlation functions for each pair of variables to be tested. To test Hypothesis 1 I cross-correlated each of the three input measures (DV, OV, and ER) with each of the technical-rational effectiveness indicators (GR, RR, and RC). I also cross-correlated each of the input measures with staff headcount (SH), and SH with each of the technical-rational effectiveness indicators. In total, to test Hypothesis 1, I calculated cross-correlation functions for 15 combinations of pairs of variables.

To test Hypothesis 2 I cross-correlated six pairs of variables, being each of the three input measures paired with each of the institutional effectiveness indicators (CV and NV). To test Hypothesis 3, I also cross-correlated six pairs of variables, these being each of the
two institutional effectiveness indicators paired with the three technical-rational effectiveness indicators. To test Hypothesis 4, I cross-correlated sick leave (SL) with each of the three technical-rational effectiveness indicators.

I was able to test all 30 of the above combinations of pairs or variables for New Zealand Police. However, because Queensland data were not available for RC, SH, or NV, I was only able to test 13 combinations for Queensland Police Service. In total I calculated 270 cross-correlation functions for New Zealand, being for 30 pairs of variables for each of the three types of pre-whitening model (ARIMA, VAR-AIC, and VAR-12) for each of the three New Zealand Police districts, and 117 cross-correlation functions for Queensland, being for 13 pairs of variables for each of the three types of pre-whitening model (ARIMA, VAR-AIC, and VAR-12) for each of the three Queensland Police Service districts. From these cross-correlation functions, I was able to characterise the relationships between variables by calculating regression coefficients for various lags to determine what, if any, statistically significant lagged relationships exist. Results are described in Chapter 7.

6.7 Step 5: Perform Granger-causality Tests

To determine if the significant lagged relationships observed in the cross-correlation functions are predictive I performed a Granger-causality test for each pair of variables. For these Granger-causality tests I formed the null hypothesis that the given $x$ variable does not Granger-cause the given $y$ variable, and the alternative hypothesis that the given $x$ variable does Granger-cause the given $y$ variable. I computed $p$-values for each test. Using this procedure I produced separate results for each type of model (ARIMA, VAR-AIC, and VAR-12).
Although statistical significance can be quantified for results of a given model, no method exists for computing the combined significance of results from different models. As discussed in Section 6.5, my purpose in using three different models to separately pre-whiten each time series is to reduce the potential for producing spurious results. I therefore treated a lag as significant if at least two of the three types of model produced similar results for both significant lags and Granger-causality tests. Having drawn conclusions about each district I then compared results across districts to see if results were consistent, and therefore generalisable.

In this chapter I described the method I used to test the four hypotheses generated from the findings of Study 1. This method involves analysing the relationship between variables identified both in Study 1 and the literature. The selected method—transfer function analysis—involves technical challenges which I addressed through aspects of the analysis design including: (a) ensuring each time series contained at least 60 data points, (b) using three different models to pre-whiten time series, and (c) using Granger-causality testing to complement cross-correlation analysis. Results from the cross-correlation analysis and Granger-causality tests are presented in Chapter 7.
Chapter 7: Study 2 Findings

7.1 Introduction

Although extensive literature exists on performance measurement in police organisations, prior research does not provide a clear picture of whether performance information influences police effectiveness. Scholars have theorised why police measure their performance and what impact performance measurement has on both the police organisation and its performance outcomes. However, these impacts have not been adequately tested, perhaps because such testing is not easy to do. In this thesis I test these impacts.

Guided by the literature reviewed in Chapters 2 and 3, I undertook an exploratory study of the perspectives and behaviours of police personnel in New Zealand and Queensland. From Study 1 I developed a conceptual model of the theoretical relationships between factors influencing the impact of performance measurement on police effectiveness. I generated four testable hypotheses based on this conceptual model:

H1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.

H2: Police do not respond to performance measures in a way that influences their organisational legitimacy.

H3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.
H4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

In Chapter 6 I described an analytical methodology to test these hypotheses. This methodology involves technical challenges which I addressed through three mechanisms: (a) ensuring each time series contained at least 60 data points, (b) using three different models to pre-whiten time series, and (c) using Granger-causality testing to complement cross-correlation analysis. In this Chapter I present results of applying this analytical methodology to each of the four hypotheses.

7.2 Descriptions of Hypotheses

7.2.1 Hypothesis 1: Performance measurement and police effectiveness

Findings of Study 1 suggest that police managers make only limited use of quantitative performance information even though performance measurement frameworks routinely report quantitative measures. These findings generated the following hypothesis: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service. To test this hypothesis I separately analysed the relationship between each of three performance measures that police monitor and three technical-rational indicators of police effectiveness in responding to calls for service. These relationships are illustrated in Figure 7.1.
Interviewees in Study 1 also perceived one aim of performance measurement as being to influence staff headcount (SH) to ensure sufficient personnel are employed to cope with demand. So, to test whether performance measures monitored by police influence staff headcount, and whether staff headcount influences police effectiveness in responding to calls for service, I test the relationships between SH and each of the variables in Figure 7.1. These relationships are illustrated in Figure 7.2.
7.2.2 Hypothesis 2: Performance measurement and institutional effectiveness

The findings from Study 1 and the literature reviewed in Chapter 2 provide evidence that police strive to maintain organisational legitimacy and avoid public criticism. Police are institutionalised organisations that use performance measurement frameworks to enhance their organisational legitimacy. However, it is unclear whether police actually respond to performance measures in a way that influences their organisational legitimacy, rather than only operate the mechanism of performance measurement to appear progressive and successful (Meyer & Rowan, 1977; Taylor, 2009; Willis et al., 2007).

From these findings I generated the following hypothesis: Police do not respond to performance measures in a way that influences their organisational legitimacy.
test this hypothesis I separately analyse the relationship between each of three measures that police monitor and two indicators of police legitimacy: complaint volumes (CV) and noise (NV). These measures and indicators were all identified in Study 1. The relationships I test in Study 2 are illustrated in Figure 7.3.

**Figure 7.3**

**Relationships Between Performance Measures and Institutional Effectiveness**

7.2.3 **Hypothesis 3: Institutional effectiveness and technical-rational effectiveness**

In Study 1 evidence was found of police managers monitoring complaints and noise in the media and parliament to determine whether police are performing effectively. If these managers act in response to variations in levels of complaints and noise, such action may influence police effectiveness in technical-rational terms. However, in Study 1 I found that police managers make only limited use of quantitative performance information. I
therefore hypothesise that variations over time in volumes of complaints and noise do not influence measures of police effectiveness in responding to calls for service.

To test this hypothesis I separately analyse the relationship between each of two indicators of organisational legitimacy—complaint volumes (CV) and noise (NV)—and three indicators of police effectiveness in responding to calls for service in technical-rational terms—general response time performance (GR), repeat calls for service (RR), and resolution rate (RR). These indicators were all identified in Study 1. The relationships I test in Study 2 are illustrated in Figure 7.4.

Figure 7.4
Relationships Between Indicators of Institutional Effectiveness and Technical-rational Effectiveness

7.2.4 Hypothesis 4: Morale and performance

In Study 1 I found that police personnel who are disillusioned with management and/or performance measurement still maintain a commitment to working hard and doing a good
job in their own eyes or in the eyes of their work-group. In Study 2, I test whether the perspectives revealed in Study 1 translate to behaviour that maintains police effectiveness in responding to calls for service even when employees are disillusioned with management. Using the amount of sick leave taken by employees as a proxy indicator for staff morale, I hypothesise that police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

To test this hypothesis I separately analyse the relationship between the number of sick days taken by police personnel in each district and indicators of police effectiveness in responding to calls for service. These measures were all identified in Study 1. The relationships I test in Study 2 are illustrated in Figure 7.5.

Figure 7.5
Relationships Between Sick Leave and Effectiveness
7.3 Interpreting Results of Hypothesis Tests

Appendix C contains the results of tests for each pair of variables identified by arrows in the diagrams in Section 7.2. In this section I describe criteria I apply to interpret these results. Then, in Section 7.4., I present findings based on these criteria.

**Results for a single district**

I produced distinct results for each of the three districts in both New Zealand Police and Queensland Police Service. As described in Chapter 6, to reduce the potential for producing spurious results, I used three different pre-whitening models. As discussed in Section 6.7, although statistical significance of results can be quantified for each model, no method exists for computing the combined significance of results from different models. It is therefore necessary to form some arbitrary criteria to interpret the overall results of different models.

As discussed in Section 6.3, my purpose in using three different models to separately pre-whiten each time series is to reduce the potential for producing spurious results. I therefore interpret a lag as significant if at least two of the three types of model produce similar results for both significant lags and Granger-causality tests. Specifically, I interpret a relationship as significant where a common statistically significant (p=0.5) cross-correlation function lag exists in at least two models and where the p-value for Granger-causality testing is less than 0.05 for both models. I interpret a result as indeterminate where a common statistically significant (p=0.5) cross-correlation function lag exists in at least two models and where the p-value for Granger-causality testing is less than 0.1 in at least one of the models. Where these criteria are not met, I interpret the result as failing to find evidence of a relationship between the variables.
Generalising results across district

Having drawn conclusions about each district, I then compare results across districts to see if results are consistent and therefore generalisable. Again, because no method exists for computing the combined significance of results from different models, it is necessary to form some arbitrary criteria to interpret the combined results of different models. For each police agency, with three models for each district, it is possible to form criteria based on results for nine models for each pair of variables.

A particular relationship may exist in one district but not others. This difference may occur because of a difference between that district and other districts. Therefore, to generalise results across districts, a similar relationship must exist in more than one district. I interpret a generalised relationship as significant where either (a) the regression coefficient for a given lag in a cross-correlation function is statistically significant (p=0.5) in at least one model from all three districts, and where the p-value for Granger-causality testing is less than 0.05 for each of these models, or (b) the regression coefficient for a given lag in a cross-correlation function is statistically significant (p=0.5) in at least two models in each of two districts, and where the p-value for Granger-causality testing is less than 0.05 for each of these models. I interpret a generalised relationship as indeterminate where a common statistically significant (p=0.5) cross-correlation function lag exists in at least three models spread across at least two districts, and where the p-value for Granger-causality testing is less than 0.1 for each of these models. Where these criteria are not met, I interpret the result as failing to find evidence of a generalisable relationship.
7.4 Findings

Appendix C contains results of statistical tests for the pairs of variables tested in Study 2. In this section I present findings from applying the criteria described in Section 7.3 to these results. Specifically, I identify whether results of transfer function analysis find evidence that the relationship between each pair of variables exists. I classify results into three categories: (a) evidence exists of a relationship, (b) I failed to find evidence of a relationship, or (c) results are indeterminate. Where I found evidence of a relationship or results are indeterminate I identify whether the relationship is reinforcing or inverse, and with what time lag. I use these findings to draw conclusions about each hypothesis.

In the tables in this section, a dash ‘-’ indicates I failed to find evidence of a relationship. An integer ‘L’ preceded by a plus ‘+’ or minus ‘-’ sign indicates I found evidence of a relationship with a lag of L months and whether the relationship is reinforcing (+) or inverse (-). An integer and sign enclosed in parentheses indicates that the relationship is indeterminate. To reduce the potential for spurious results cross-correlations for lags greater than 6 months are ignored. This is because the standard error estimates of regression coefficients for lags greater than approximately 5 or 6 can be biased, and this may affect the significance tests of the correlation coefficients (Akarca & Long Li, 1979; Erol & Yu, 1987; Haugh, 1976).
7.4.1 Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service

New Zealand districts

Table 7.1 presents the results of applying the criteria described in Section 7.3 to the statistical tests I conducted on New Zealand data to test Hypothesis 1. The majority of tests for each district failed to find evidence of a relationship between the variables. In District 2, offence volume (OV) appears to influence resolution rate (RR) at lag = 1 (i.e. an increase in the number of offences appears to produce an increase in resolution rate one month later\(^{15}\)). Also an increase in the number of calls for service (DV) and/or emergency response-time performance (ER) appears to produce an increase in volume of repeat calls for service (RC) two months later. An increase in the number of calls for service (DV) also appears to produce a decrease in staff headcount (SH) five months later, and in District 3 an increase in emergency response-time performance (ER) appears to produce a decrease in staff headcount (SH) one month later. All other results are either indeterminate or provide no evidence of a relationship between the variables.

\(^{15}\) And/or a decrease in the number of offences may produce a decrease in resolution rate one month later. This relationship is reinforcing rather than inverse. To aid in readability I express a reinforcing relationship as ‘an increase in x appears to produce an increase in y’; conversely, I express an inverse relationship as ‘an increase in x appears to produce a decrease in y’.
### Table 7.1
**New Zealand Results for Hypothesis 1**

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 1</th>
<th>District 2</th>
<th>District 3</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RC</td>
<td>-</td>
<td>+2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-</td>
<td>+1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-</td>
<td>-</td>
<td>(-1)</td>
<td>(-1)</td>
</tr>
<tr>
<td>ER-RC</td>
<td>-</td>
<td>+2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>(-4)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-SH</td>
<td>-</td>
<td>-5</td>
<td>(-5)</td>
<td>(-5)</td>
</tr>
<tr>
<td>OV-SH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-SH</td>
<td>-</td>
<td>-</td>
<td>-1</td>
<td>-</td>
</tr>
<tr>
<td>SH-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Queensland districts**

Table 7.2 presents the results of applying the criteria described in Section 7.3 to the tests I conducted on Queensland data to test Hypothesis 1. Repeat calls for service (RC) and Staff Headcount (SH) data were not available for Queensland, so fewer tests were able to be performed than were possible for New Zealand.
Table 7.2
Queensland Results for Hypothesis 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 4</th>
<th>District 5</th>
<th>District 6</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-GR</td>
<td>+3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-</td>
<td>-</td>
<td>+1</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>(-2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

In District 6, an increase in the number of offences (OV) appears to produce an increase in resolution rate (RR) one month later. No other tests for Districts 5 or 6 found evidence of a relationship between the variables. In contrast to Districts 5 and 6, only two of the six tests conducted on District 4 data failed to find evidence of relationships between the variables. Three tests found evidence of relationships, and the remaining test result was indeterminate.

Section 5.2.3 describes one interviewee in Study 1 as a notable exception to other managers. Whereas most interviewees make only limited use of performance information, this police manager described using quantitative performance information to identify problems, evaluate impacts of actions, motivate personnel to undertake previously unpopular tasks, and increase productivity of personnel. This interviewee is a manager in District 4, which is the only district where I find substantial evidence that performance measures monitored by police (DV, OV, & ER) influence (Granger-cause) measures of police effectiveness in responding to calls for service (GR & RR).
**Generalising findings**

If police respond to performance measures in a way that influences police effectiveness I would expect to find evidence of a relationship between the pairs of variables tested. I applied the criteria described in Section 7.3 for generalising results across districts. On this basis, for 13 of the 15 pairs of variables in New Zealand and all six of the pairs of variables in Queensland I failed to find evidence of a generalisable relationship. Results are indeterminate for the remaining two New Zealand pairs of variables (ER-GR and DV-SH). Although these findings do not prove Hypothesis 1, in no test could I reject the null hypothesis which is Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in responding to calls for service.

I did find evidence of performance information in District 4 influencing police effectiveness in responding to calls for service: Only two of the six tests on District 4 data failed to find evidence of a relationship between performance information and measures of police effectiveness. Importantly, District 4 contains the manager who in Study 1 was identified as a notable exception to other interviewees in that this manager actively uses performance information to influence decisions. Consequently, for District 4 and only District 4, Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in technical-rational terms, is rejected. These findings taken together suggest that it is possible for performance information to influence police effectiveness in responding to calls for service, but in most instances there is little evidence that it does.
7.4.2 Hypothesis 2: Police do not respond to performance measures in a way that influences their organisational legitimacy

New Zealand districts

Table 7.3 presents the results of applying the criteria described in Section 7.3 to the tests I conducted on New Zealand data to test Hypothesis 2. In District 1, results are indeterminate for whether emergency response time performance (ER) influences institutional effectiveness. In all other tests for District 1 and all tests for Districts 2 and 3 I failed to find evidence of a relationship between the variables.

Table 7.3
New Zealand Results for Hypothesis 2

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 1</th>
<th>District 2</th>
<th>District 3</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-CV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-NV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-NV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-CV</td>
<td>(-1)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-NV</td>
<td>(-2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Queensland districts

Table 7.4 present the results of applying the criteria described in Section 7.3 to the tests I conducted on Queensland data to test Hypothesis 2. Noise (NV) data were not available for Queensland, so fewer tests were able to be performed than were possible for New
Zealand. As shown in Table 7.4, I failed to find evidence of a relationship between the variables within any of the three Queensland districts.

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 4</th>
<th>District 5</th>
<th>District 6</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-CV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-CV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Generalising findings**

I applied the criteria described in Section 7.3 for generalising results across different sites. Applying these criteria failed to find any evidence of a generalised relationship across districts in either New Zealand Police or Queensland Police Service. Although these findings do not prove Hypothesis 2, in no test could I reject the null hypothesis that police do not respond to performance measures in a way that influences their organisational legitimacy.

**7.4.3 Hypothesis 3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service**

**New Zealand districts**

Table 7.5 presents the results of applying the criteria described in Section 7.3 to the tests I conducted on New Zealand data to test Hypothesis 3. An increase in noise (NV) appears
to produce a decrease in general response time performance (GR) four months later in all three districts. A graph of the NV time series is presented in Figure 7.6.

Table 7.5
New Zealand Results for Hypothesis 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 1</th>
<th>District 2</th>
<th>District 3</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV-GR</td>
<td>(+2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-GR</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>NV-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-RR</td>
<td>-</td>
<td>-</td>
<td>+1</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 7.6
Time Series of NV

Correlograms and associated test statistics for each of the three pre-whitened time series for NV are presented in Figures 7.7, 7.8, and 7.9. By definition, past values of a white
noise time series are uncorrelated with future values of the same series (McCleary & Hay, 1980; Wei, 2006). Therefore, regression coefficients of the autocorrelation (AC) and partial autocorrelation function (PAC) for all lags $> 0$ should be close to zero for a time series that resembles white noise. As shown in Figures 7.7, 7.8, and 7.9, the magnitude of regression coefficients for lags 1 and 2 are less than 0.063 in all of these models. However, regression coefficients for the partial autocorrelation function (PAC) is greater than 0.16 for lags 3 and 4 in all three models. These patterns show that, although all three models do a good job of removing autocorrelation for lags of one or two months, none of the models do a good job of removing autocorrelation for lags of three or four months. Results of cross-correlations of NV with other time series should therefore be treated with caution for apparent relationships of three or four months delay. Specifically, the apparent reduction in general response time performance (GR) produced by an increase in noise (NV) may be a spurious result.

**Figure 7.7**

**Correlogram of ARIMA Pre-whitened NV**

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>1.0</td>
<td>-0.009</td>
<td>0.0055</td>
<td>0.941</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>2</td>
<td>0.022</td>
<td>0.022</td>
<td>0.0361</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>3</td>
<td>0.166</td>
<td>0.157</td>
<td>1.8409</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>4</td>
<td>-0.195</td>
<td>-0.197</td>
<td>4.3621</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>5</td>
<td>-0.121</td>
<td>-0.138</td>
<td>5.3519</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>6</td>
<td>0.192</td>
<td>0.192</td>
<td>7.8874</td>
</tr>
</tbody>
</table>
Table 7.5 shows that in District 3, an increase in noise (NV) appears to produce an increase in resolution rate (RR) one month later. However, no evidence was found of a similar relationship in the other two districts. In District 1, results were indeterminate for whether complaint volumes (CV) influence general response time performance (GR). All other tests failed to find evidence of a relationship between the variables.

**Queensland districts**

Table 7.6 presents the results of applying the criteria described in Section 7.3 to the tests I conducted on Queensland data to test Hypothesis 3. Repeat calls for service (RC) and Noise (NV) data were not available for Queensland, so fewer tests were able to be performed than were possible for New Zealand. For the remaining two pairs of variables...
(CV-GR and CV-RR), all tests failed to find evidence of a relationship between the variables for Districts 4 and 6. Results for CV-GR in District 5 are indeterminate.

Table 7.6
Queensland Results for Hypothesis 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 4</th>
<th>District 5</th>
<th>District 6</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>(+2)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Generalising findings

I applied the criteria described in Section 7.3 for generalising results across different sites. Applying these criteria failed to find any evidence of a generalised relationship across districts in either New Zealand Police or Queensland Police Service apart from the relationship between noise (NV) and general response time performance (GR). As discussed above, this result for NV should be treated with caution because none of the pre-whitening models for NV adequately removed serial correlation at lags 3 and 4. I therefore ignore the apparent reduction in general response time performance (GR) produced by an increase in noise (NV). Although these findings do not prove Hypothesis 3, in no test could I reject the null hypothesis that variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.
7.4.4 Hypothesis 4: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take

Hypotheses 1, 2, and 3 all consider potential causal relationships, so do not require testing simultaneity (relationships for lag 0—no delay). However, Hypothesis 4 simply explores whether sick leave is related to measures of police effectiveness in responding to calls for service. Therefore, the relationships between variables at lag = 0 also need to be tested. Granger-causality testing involves testing if past values of one variable help to predict future values of another, so Granger-causality testing is not applicable for testing relationships for lag = 0. However, cross-correlation is still possible. I therefore cross-correlate sick leave (SL) with general response time performance (GR), repeat calls for service (RC), and resolution rate (RR). As indicated in Appendix C, only one of the 45 cross-correlation functions gives a correlation coefficient at lag 0 of greater than two standard errors. Therefore, overall, I failed to find evidence that indicators of police effectiveness are sensitive to variations in staff morale, as indicated by the amount of sick leave that employees take.

I also tested whether staff morale has a delayed effect on police effectiveness in responding to calls for service. Tables 7.7 and 7.8 present results for lags > 0 (i.e. delays of one month or more) for New Zealand and Queensland respectively. Repeat calls for service (RC) data were not available for Queensland, so fewer tests were able to be performed than were possible for New Zealand.
Table 7.7
New Zealand Results for Hypothesis 4

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 1</th>
<th>District 2</th>
<th>District 3</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>-</td>
<td>-3</td>
<td>(-2)</td>
<td>-3</td>
</tr>
<tr>
<td>SL-RC</td>
<td>-</td>
<td>+4</td>
<td>-1</td>
<td>-</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7.8
Queensland Results for Hypothesis 4

<table>
<thead>
<tr>
<th>Variables</th>
<th>District 4</th>
<th>District 5</th>
<th>District 6</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>-</td>
<td>-</td>
<td>(-5)</td>
<td>-</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**New Zealand results**

In District 2, an increase in sick leave (SL) appears to produce a decrease in general response time performance (GR) after three months, and in District 3 results are indeterminate as to whether SL reduces GR after two months. In considering whether any consistent relationships exist across different districts, I note that one model for District 1 and two models for District 2 produce a statistically significant negative correlation and Granger causality test (p<0.05) at lag = 3 (see Appendix C), which suggests that an increase in sick leave appears to decrease general response time performance three months later. One model for District 3 produces a statistically significant negative correlation and Granger causality test (p=0.0482) at lag = 2, and another model produces a similar pattern but slight weaker Granger causality test result (p=0.0773). So, taken together these results indicate that an increase in sick leave (SL) appears to produce a
decrease in general response time performance (GR) two or three months later. It may be that police maintain their performance in responding rapidly to calls for service when their morale is low, but that low morale has a delayed detrimental influence on police effectiveness.

Results for the relationship between sick leave (SL) and repeat calls for service (RC) vary between districts. Tests for District 1 failed to find evidence of a relationship between these variables. However, in District 2 an increase in sick leave (SL) appears to produce an increase in repeat calls for service (RC) four months later, and in District 3 an increase in SL appears to produce a decrease in RC one month later. Therefore, there may be some interaction between these variables, but no consistent pattern exists. I failed to find evidence of a relationship between sick leave (SL) and resolution rate (RR) in any district.

Queensland results

Results are indeterminate for whether sick leave (SL) produces a decrease in general response-time performance (GR) in District 6. All other tests failed to find evidence of relationships between variables. Therefore, overall I failed to find evidence of a generalizable relationship in Queensland between sick leave (SL) and measures of police effectiveness in responding to calls for service.

7.5 Summary of Findings

This chapter presented results of Study 2. In sum, the findings of Study 2 help to address the core aims of this thesis by testing the four hypotheses that arose from Study 1. In Study 2 I failed to find evidence that the monitoring of performance measures by police influences police effectiveness in responding to calls for service in general. However,
performance monitoring may have an influence in isolated instances: I found some evidence in one district of performance information influencing police effectiveness. Importantly, I observe that this is the same district in which Study 1 found a manager who is a notable exception to other managers in using performance information effectively. Consequently, for this district and only this district, Hypothesis 1: The monitoring of performance measures by police does not influence police effectiveness in technical-rational terms, is rejected. Applying the criteria described in Section 7.3 for generalising results across districts, in none of the tests could I reject Hypothesis 1. Therefore, taken together with the findings of Study 1, I conclude that it is possible for performance information to influence police effectiveness in responding to calls for service, but in most instances there is little evidence that it does.

In Study 1 I found numerous examples of police monitoring indicators of organisational legitimacy and acting to enhance their organisational legitimacy. In Study 2 I examined the relationships between performance measures monitored by police and indicators of organisational legitimacy, and failed to find evidence of a relationship. Consequently, I was unable to reject Hypothesis 2: Police do not respond to performance measures in a way that influences their organisational legitimacy. I also examined the relationships between these indicators of organisational legitimacy and technical-rational indicators of police effectiveness in responding to calls for service, and failed to find evidence of a relationship. Consequently, I was unable to reject Hypothesis 3: Variations over time in volumes of complaints and noise do not influence indicators of police effectiveness in responding to calls for service.
Finally, I tested a fourth hypothesis: Police effectiveness in responding to calls for service is insensitive to variations in staff morale, as indicated by the amount of sick leave that employees take. I failed to find evidence of morale having an immediate influence on police effectiveness in responding to calls for service. However, I found some evidence for a delayed effect. This delayed relationship is more evident in New Zealand than in Queensland, and varies for different dimensions of police effectiveness. For example, I failed to find a relationship between morale and resolution rate, but I found some evidence of low morale having a delayed detrimental effect on general response time performance.

In Chapter 8 I discuss these findings in more detail. In particular, I discuss their theoretical implications, and the implications for policy and practice. I also discuss the limitations of these findings, and suggest implications for future research.
Chapter 8: Implications for Theory, Policy, and Practice

8.1 Introduction

Police play an important role in maintaining a functioning society. Democratic state governments utilise police to maintain their legitimacy with constituents. They do this by requiring that police demonstrate authoritative intervention to restore order when required (Bayley, 1994b) and demonstrate symbolic justice, showing offenders and the public that a regime of law exists (Silberman, 1978). Democratic state governments wish to reinforce their legitimacy as a democracy by demonstrating responsiveness to citizens (Herbert, 2006). Police help to provide this responsiveness by providing citizens with a sense of personal safety engendered by confidence that someone who can prevent harm can be called on in an emergency or in a dangerous situation. Police seek to maintain and enhance their own organisational legitimacy, both with the public and government, thereby securing status and resources (Herbert, 2006; Lawton et al., 2000; Tyler & Huo, 2002).

Demonstrating responsiveness to calls for service from citizens requires that police manage numerous challenges and pressures (Corboy et al., 2005; Gunal et al., 2007; Kuhn & Hoey, 1987; Moskos, 2007; Sacks et al., 1993; Sparrow et al., 1990). To manage these challenges police introduced performance measurement frameworks. Performance measurement can help manage complexity by enabling members of an organisation to make sense of what is happening in complex organisational contexts (Gill, 2011). However, some scholars question whether police managers have skills to use performance information effectively (Andersson & Tengblad, 2009; Coleman, 2008); others suggest police organisations use performance measurement frameworks as legitimising structures.
to appear legitimate, rather than to improve service delivery to citizens (Taylor, 2009; Willis et al., 2007).

My thesis explores the influence of performance measurement on the effectiveness of police in responding to calls for service from citizens. I studied the perspectives and behaviours of police personnel involved in delivering or managing the service of responding to calls for service. In doing so I created new insights into values and cultural norms within police and how these influence how police manage their performance. These insights generated hypotheses that I tested to better understand the influence of performance measurement on the effectiveness of police in responding to calls for service from citizens. The findings of this thesis make an important contribution to knowledge about the use and impact of performance information in police organisations.

In this chapter I revisit the key findings of this thesis and discuss their implications. In Section 8.2 I outline the key findings of Study 2, and contextualise these with evidence from Study 1. Here I determine what influence performance measurement has and—perhaps more importantly—does not have on the effectiveness of police in responding to calls for service from citizens. In Section 8.3 I identify implications for police policy and practice. Although this thesis provides insights that help to answer my research questions, some questions remain unanswered. In section 8.4 I discuss some caveats and limitations to my research and recommend avenues for future research that may extend the understanding provided by this thesis.
8.2 Main Findings

Central to my research are four main aims:

Aim 1. examine what police understand effectiveness in responding to calls for service to be;

Aim 2. determine what influence performance measurement has on police effectiveness;

Aim 3. explore why performance measurement has the influence it does on police effectiveness; and

Aim 4. explore the influence of performance measurement on the perspectives and behaviours of police personnel.

I commenced my research by reviewing relevant literature, and identified gaps in knowledge provided by prior research. To address these gaps I formed three main research questions:

RQ 1. How do police managers use performance information, and do they use performance information effectively?

RQ 2. Does performance measurement help or hinder police effectiveness in either technical-rational or institutional terms?
RQ 3. What influence does performance measurement in police organisations have on loose coupling, and what influence does loose coupling have on police effectiveness?

To answer these research questions I undertook two interconnected studies which build on prior research reviewed in Chapters 2 and 3. In this section I restate my findings from Studies 1 and 2.

In Chapter 2 I showed that performance measurement—introduced as part of NPM reforms—can and has delivered performance improvements in public sector organisations (de Boer & Evans, 1996; Gill, 2011; Pollitt & Bouckaert, 2011; Toime & Steele, 1991). However, police failed to adopt NPM reforms as effectively as some other public sector agencies (Ashby et al., 2007; Cope et al., 1997; Legrand & Bronitt, 2012; M. H. Moore & Braga, 2003). Such reforms challenged the values of police (Lapsley, 2008; M. H. Moore & Braga, 2003); police were able to leverage political pressures to shield themselves from the full effect of NPM reforms (Cope et al., 1997; Savage, 2007). Prior research found evidence of loose coupling within police organisations in which employees become decoupled from structures and policies set by senior management (Collier, 2001; Lawton et al., 2000; Meyer & Rowan, 1977). However, prior research fails to paint a clear picture about the relationship between performance measurement and police effectiveness.

**8.2.1 Breadth of performance measures**

Literature reviewed in Chapter 2 suggests that police closely monitor measures related to crime but ignore the many other things that police do (Fleming, 2009; Kelling, 1992;
Legrand & Bronitt, 2012; Maguire, 2003; Mastrofski, 2004; M. H. Moore & Braga, 2003; Skogan & Frydl, 2004). These scholars argue that the absence of broader performance measures prevents police from being as effective as they could be. In Study 1, I found that police personnel involved in delivering or managing a service of responding to calls for service from citizens identify a broad range of dimensions of police effectiveness. The narratives from Study 1 do identify performance measures related to crime, such as offence volumes and crime resolution rates. However, interviewees also identified measures that relate to both crime and non-crime activities of police, such as response timeliness, complaints, customer satisfaction, noise in the media or parliament, volumes of calls for service, and volumes of repeat calls for service to the same address (see Section 5.3.4).

Although police personnel may be able to identify a broad range of performance measures this is insufficient to ensure performance information influences police effectiveness; performance information also has to be used (de Lancer Julnes & Holzer, 2001; Van Dooren & Van de Walle, 2008). In Study 1 I identified a broad range of performance measures, but found little evidence of police managers using performance measures to influence their decisions. Consistent with the findings of (Willis, 2011), the most common purpose identified in Study 1 of performance measurement is to justify or plan the number of staff to deploy. Performance measurement is rarely used to promote innovative responses to problems or to analyse the outcomes of strategies and activities to determine which are effective. Instead, the dominant approach used for managing performance involves checking that the work of subordinates was appropriate or complied with standard practice (see Section 5.2). Some scholars suggest police managers may lack the skills to use performance information effectively (Andersson & Tengblad, 2009;
Coleman, 2008). Based on these suggestions and findings of Study 1, in Study 2 I tested whether the monitoring of performance measures by police influences police effectiveness in responding to calls for service.

8.2.2 The influence of performance measurement on police effectiveness

Prior research shows that the introduction of OPRs influences police performance outcomes (Chilvers & Weatherburn, 2004; Mazerolle et al., 2007). However, this impact on outcomes may be due to police commanders reacting to the announcement that the OPR was imminent. There is a gap in the literature on whether performance information monitored through OPRs or routine performance reporting influences police effectiveness. In this thesis I address this gap by identifying performance measures monitored in police performance frameworks then testing whether variations in these measures influence measures of police effectiveness.

Findings of Study 1 suggest that police do not respond to performance measures in a way that influences their service delivery performance (Hypothesis 1) or organisational legitimacy (Hypothesis 2). The results of Study 1 also suggests that, although police managers monitor indicators of organisational legitimacy, variations in these indicators over time do not appear to influence police service delivery performance (Hypothesis 3). In Study 2 I analysed time series data to test whether these hypotheses could be rejected. I failed to find sufficient evidence in the data to reject any of these hypotheses suggested in Study 1. These findings taken together suggest that, in general, police managers do not respond to quantitative measures in a way that influences their effectiveness in either
technical-rational (such as service delivery performance) or institutional (such as organisational legitimacy) terms.

However, performance monitoring may have an influence in isolated instances: In Study 2 I found some evidence in one district of performance information influencing police effectiveness. I observe that this is the same district in which Study 1 found a manager who is a notable exception to other managers in using performance information effectively. This finding suggests that the monitoring of performance measures by police can influence police effectiveness in responding to calls for service if police managers are capable of using performance information effectively. However, I found little evidence that most police managers demonstrate such capability.

8.2.3 Why police measure performance

If police do not use performance measurement to improve their effectiveness this raises the question of what motivates police to use performance measurement. Taylor (2009) and Willis et al. (2007) argue that police use performance measurement frameworks to gain legitimacy and appear progressive, rather than to improve performance. In studying the impact of OPRs on police performance Chilvers and Weatherburn (2004) and Mazerolle et al. (2007) show that performance improves following the introduction of OPRs. However, this improvement is not necessarily caused by police managers responding to data reported in the OPRs, but could result from police managers acting to mitigate the threat to their legitimacy that an OPR creates. Mazerolle et al. (2007) acknowledge the possibility that the improvement in performance may result from an “announcement effect” (p. 242) whereby police managers alter their management and operational strategies in advance of the OPR occurring. The improvements in
performance found by Chilvers and Weatherburn (2004) and Mazerolle et al. (2007) may therefore be caused by the desire of police managers to appear to be performing well, rather than by the insights gained through the performance information being monitored.

In Study 1 I found evidence of police using performance reporting to look good, justify resources, and defend themselves from criticism (see Section 5.3.1). Institutional theory gives an explanation for this finding: Institutionalised organisations strive to maintain their legitimacy to secure resources and survive (Meyer & Rowan, 1977; R. W. Scott, 2004; Suchman & Edelman, 1996). By adopting performance measurement frameworks police gain legitimacy by demonstrating structural conformity (isomorphism) with other organisations (DiMaggio & Powell, 1983). Given that satisfaction with police is uncorrelated with actual service quality (Brown & Coulter, 1983; Stipak, 1979) it is unsurprising that police managers focus on managing organisational legitimacy directly, rather than only managing service delivery performance.

8.2.4 Loose coupling

In Study 1 I found that junior police personnel more frequently identify technical-rational goals, such as apprehending offenders and preventing harm, than do senior managers. However, senior managers more frequently identify objectives that support institutional goals such as organisational legitimacy and demonstrating respect for the caller (see Section 5.4.1). Lawton et al. (2000) and Meyer and Rowan (1977) argue that perceptions by an organisation’s employees of such misalignment or confusion between technical-rational goals and institutional goals can lead to loose coupling. Brunetto and Farr-Wharton (2005) argue that “successful implementation of policies requires that senior and lower managers must be in congruence in relation to the stated goals and objectives of a
new policy” (p. 221). Therefore the loose coupling found in Study 1 can threaten the integrity of performance measurement frameworks, and thereby threaten police effectiveness (Andersson & Tengblad, 2009; Brunetto & Farr-Wharton, 2003, 2005; Lawton et al., 2000; Lipsky, 1980; McKevitt & Lawton, 1996). In Study 1 I found that, rather than looking to management to ensure their work groups are performing well, work groups appear to self-regulate the behaviour of their members through a culture of strong peer pressure to work hard for the team. So, although this culture drives commitment by individuals to perform, this finding suggests that formal performance measurement frameworks in police may have little influence on the effectiveness of police.

Loose coupling may threaten the effectiveness of performance measurement frameworks (Andersson & Tengblad, 2009; Brunetto & Farr-Wharton, 2003, 2005; Lawton et al., 2000; Lipsky, 1980; McKevitt & Lawton, 1996). However, the culture of strong commitment by police personnel to perform may buffer the organisation from the effects of staff disillusionment with police management (see Section 5.4.2). In Study 2 I failed to find evidence that police effectiveness in responding to calls for service is affected by variations in staff morale—at least in the short term. Police in New Zealand and Queensland maintain levels of performance in apprehending offenders and responding to calls for service in a timely way irrespective of variations in employee morale.

8.2.5 Conclusions

Although police in New Zealand and Queensland monitor measures relevant to fighting crime, police personnel involved in delivering or managing a service of responding to calls for service have a good understanding of broader dimensions of police effectiveness. This understanding is reflected in performance measures that include a strong focus on
crime, but also include broader measures relevant to satisfying citizens and enhancing police organisational legitimacy with citizens. However, although performance measurement can influence police effectiveness in responding to calls for service, in general, the practice of measuring performance appears to have little influence in these two police organisations. In fact, staff disillusionment with performance management frameworks may contribute to loose coupling between first line supervisors and more senior managers.

Although this loose coupling can harm the effectiveness of performance measurement (Brunetto & Farr-Wharton, 2005; Lawton et al., 2000), it may also buffer police personnel from institutional pressures that concern senior police managers by "maintain[ing] standardized, legitimating, formal structures while their activities vary in response to practical considerations" (Meyer & Rowan, 1977, p. 357). Police personnel appear to have a strong work ethic to do a good job for their team and deliver a good service to citizens (see Section 5.4.2). This work ethic, together with the buffering effect provided by first line supervisors, may explain why police performance appeared unaffected by variations in staff morale (Hypothesis 4).

Performance measurement can have a positive influence (Gill, 2011; O'Flynn, 2007), as demonstrated by District 4. However, this requires a different approach than most police managers take. In Section 8.3 I discuss this and other implications for policy and practice.

**8.3 Implications for Policy and Practice**

The findings summarised in Section 8.2 create a dichotomy for performance measurement in police. On one hand performance measurement as currently practised does not appear
to significantly influence police effectiveness in responding to calls for service. However, in spite of this, personnel who deliver the service maintain a commitment to doing a good job, and performance measurement frameworks may serve a purpose of helping to legitimise police organisations. Therefore, maintaining the current state has merit in that performance measurement acts as a buffer that allows committed police personnel to respond to calls for service as they deem appropriate.

On the other hand, delivering an effective service of responding to calls for service involves numerous complex challenges (Gunal et al., 2007; Klinger & Bridges, 1997; Moskos, 2007). Not all police personnel who deliver the service are satisfied that police deliver an efficient and effective service (see Chapter 5). Performance measurement can help organisations to make sense of what is happening in complex organisational contexts (Gill, 2011), and even though performance measurement does not always influence decision making it can do so (Taylor, 2009). Therefore, because police executives have been entrusted with the responsibility for deploying the police resource on behalf of their stakeholders, using performance measurement to influence decisions is the ethically and morally right thing to do (M. H. Moore & Braga, 2003). In this section I discuss the arguments for each of these perspectives.

8.3.1 Argument for preserving the current state

I characterise a cultural norm in police agencies in which senior managers respond to institutional pressures to maintain organisational legitimacy. These managers operate performance measurement frameworks which act as a buffer that allows committed police personnel to deliver a service of responding to calls for service as they deem appropriate. Although these frameworks include reporting of quantitative performance information
most police managers do not make effective use of this information (see Section 5.2). This may be because they lack the skills to do so (Coleman, 2008), or because they have been so socialised into thinking like police officers that they are unable to think like managers (Andersson & Tengblad, 2009).

Van Dooren et al. (2010) suggest that barriers to using performance information exist at three levels: (a) individual psychological barriers such as limited cognitive ability, (b) cultural barriers within an organisation, and (c) institutional barriers created by frameworks used by professions which influence how choice is framed and sense-making is shaped. Police are highly institutionalised (Crank, 2003) loosely coupled (Willis et al., 2007) organisations whose management is dominated by members of a single profession (Andersson & Tengblad, 2009; Jacobs et al., 2014) who do not make effective use of performance information (see Section 5.2). Therefore all three of these barriers exist in police.

To replace police managers with people who have different experience and skills from police officers may reduce the barriers to using performance information effectively. However, implementing such a reform may reduce police legitimacy with citizens and may be successfully challenged by police unions, as occurred in Britain in the 1990s (Ashby et al., 2007; Cope et al., 1997). Therefore, because the barriers to change exist, and because police personnel are committed to doing a good job, police managers provide a useful function by simply striving to maintain organisational legitimacy. Maintaining organisational legitimacy provides a buffer that allows police personnel to focus on delivering service to citizens.
8.3.2 Argument for changing the current state

By failing to use performance information effectively, police are less effective than they could be in responding to calls for service. In the extreme police may fail to save lives, as illustrated by the deaths of Iraena Asher and Cherie Cundy (see Chapter 1). Some police personnel interviewed in Study 1 would like police to use performance measurement frameworks to avoid such crises. This perspective is clearly illustrated in a statement made by a performance manager interviewed in Study 1:

One of the great indicators of performance is when things go really badly wrong – way too late by then. Priority Policing \(^\text{16}\) is a classic example of that. We should never have had a need for that you know, but that's one of the things police do, they wait for things to go bad before that happens. [QM007]

Even if performance information does not influence decisions of police managers, the practice of reporting a given performance measure can shape the perspectives that police personnel have of what effective performance looks like. This is illustrated by the contrast between Queensland Police Service and New Zealand Police. Queensland Police Service produces a routine performance report on repeat calls for service; New Zealand Police does not. Twenty nine percent of Queensland interviewees in Study 1 identified repeat calls for service as a measure of performance, compared with only two percent of New Zealand interviewees.

\(^\text{16}\) Priority Policing is a policy introduced by Queensland Police Service following a heavily publicised failure to respond to a call for service in time to prevent loss of life.
If police become able to translate awareness of what effective performance looks like into action that brings about effective performance in terms that police can recognise, then they may be able to use performance measurement to improve their effectiveness. Studies 1 and 2 found evidence in one location (District 4) of performance information being used effectively to influence performance. However, this instance was an exception.

It may be possible to reduce the barriers that exist in police to using performance information. Organisations with analytical capability more readily use performance information (Bourdeaux & Chikoto, 2008; Dull, 2009). Skogan and Frydl (2003) describe improvements in policing in the USA that were influenced by people not necessarily employed by the police evaluating whether police are effective. Speklé and Verbeeten (2014) recently found that “exploratory use of the performance measurement system tends to enhance performance” (p. 131). Bayley (2008) argues that U.S. police have improved their effectiveness by creating openness to criticism of the police establishment (p. 15). Therefore, by employing people who have analytical skills in influential roles, and by increasing external scrutiny, police may be able to make more effective use of performance information.

**8.4 Caveats and Opportunities for Future Research**

In this thesis I find that the monitoring of performance measures by police appears to have little influence of police effectiveness in responding to calls for service. Police managers strive to maintain organisational legitimacy, and more readily manage performance by checking the work of subordinates than by responding to patterns in performance indicators. Perceptions by police personnel of a lack of effectiveness of performance measurement contribute to police being loosely coupled organisations. Junior police
personnel become decoupled (Meyer & Rowan, 1977, p. 357) from organisational frameworks imposed by senior management. Instead of formal performance frameworks driving performance, peer pressure within workgroups encourages police personnel to maintain a commitment to doing a good job.

My findings shed light on key issues concerning the influence of performance measurement on the effectiveness of police in responding to calls for service. However, like all research, these findings must be considered in the context of the limitations of the studies. In this section I outline caveats to my findings, and suggest directions for future research.

### 8.4.1 Perspectives of citizens

Although the literature I reviewed in Chapter 3 includes prior research on perspectives of citizens, the studies I undertook as part of this thesis explored only the perspectives of police personnel, then tested hypotheses about the relationships between variables identified by these police personnel. Study 1 found that police personnel in New Zealand and Queensland have a good understanding of dimensions of police effectiveness identified in the literature. However, it may be that citizens in New Zealand and Queensland have different expectations of police from the dimensions identified in the literature and in Study 1.

If police are unaware of citizens’ expectations then police may fail to meet the needs of citizens. Before concluding that police have a good understanding of all dimensions of police effectiveness, more research should be conducted to determine the extent to which citizens’ expectations of police align with the perspectives of police personnel on what
constitutes effective police performance in responding to calls for service from citizens. One particular dimension worth exploring is the extent to which citizens expect police to attend the scene of the callout. Study 1 identified a difference between perspectives of police personnel in New Zealand and Queensland. Queensland personnel more readily identified the importance of attending the scene than did personnel from New Zealand Police. It is unclear whether this difference reflects different expectations citizens have or different structures or beliefs within the two police agencies studied.

Because of recent structural changes in both New Zealand Police and Queensland Police Service it may be difficult to fully answer this question. At the time I conducted Study 1 New Zealand Police operated a more centralised structure than Queensland Police Service. New Zealand Police had only three PCCs, and was in the process of implementing a PCC to handle non-emergency calls. Queensland Police Service operated numerous smaller PCCs of varying sizes, and was also in the process of implementing a PCC for non-emergency calls. These structural changes may affect the perceptions police personnel have about the importance of attending the scene in person. However, irrespective of the way police structure themselves, citizens may still have different expectations of police than police have of themselves.

8.4.2 Limitations of monthly time series data

This thesis provides insight into the relationship between performance measurement and police effectiveness. New Zealand Police and Queensland Police Service supplied monthly time series data that enabled me to assess whether measures that are monitored in a given month can help predict changes in other variables. Analysis of monthly time

series data is appropriate for understanding the influence of monthly routine reporting cycles where police managers are given the opportunity each month to make decisions that affect values of variables in future months. However, monthly time series data are insufficient to assess more immediate relationships. For example, I was able to analyse whether measures monitored by New Zealand Police in a given month influence decisions about changes in overall staffing levels for a given district in subsequent months. However, I was unable to analyse whether police reallocate staff between functions within districts based on week-to-week, day-to-day, or even hour-to-hour variations in demand.

8.4.3 Limitations of transfer function analysis

I used a transfer function analysis approach (Schwert, 1979; Wei, 2006) in Study 2 to test the relationship between variables. Transfer function analysis methodologies are prone to producing spurious results (Ashley et al., 1980; Erol & Yu, 1987; Schwert, 1979). To address the potential for obtaining spurious correlations I used the double pre-whitening procedure proposed by Haugh (1976) to transform time series into an approximate white noise form. In reviewing a range of alternative methods Pierce and Haugh (1977) find that the double pre-whitening procedure is be more robust than other methods they reviewed.

The predictive power (ability to reject Type II errors) can be increased, and the potential for spurious correlations (Type I errors) can be reduced, by lengthening the time series. Although 50 data points may be sufficient to produce useful results, 100 or even 200 data points provide greater robustness (Haugh, 1976; McCleary & Hay, 1980). Time series used in Study 2 ranged in length from 60 to 82 data points. Sixty monthly data points provide a time series covering a five year period. Lengthening these series was not
practical because consistent data spanning a longer period did not exist due to changes in
the police ICT systems from which data were sourced. Consequently, it was not possible
in Study 2 to eliminate the potential for spurious correlation. To be confident that the
findings of Study 2 are robust, future research should attempt to replicate this study, if
possible, using longer time series. Replicating Study 2 would require different data sets
be sourced from different systems, most likely in different police agencies. The question
of generalisability therefore arises.

8.4.4 Generalisability beyond Queensland and New Zealand

A final limitation concerns the generalisability of my findings. My research extends
criminological knowledge about the relationship between how police manage their
performance and how effective they are in responding to calls for service from citizens.
To date no prior study has examined the association between performance measures and
police effectiveness. My studies examine these associations in New Zealand and
Queensland. They produce broadly consistent results across both agencies. By producing
consistent results across two different police agencies in two different countries the
generalisability of my findings are strengthened. However, as discussed in Chapter 4,
Australia and New Zealand share a similar culture and have similar police functions.
Although the premise for my studies arose from the literature on police in numerous
different countries, this literature—particularly that reviewed in Section 2.4—identifies
differences in the structures and performance frameworks used in different countries. It
follows that findings for Queensland and New Zealand may differ from findings of similar
studies that might be conducted in other countries. To generalise findings from this thesis,
further similar studies should be conducted in other countries.
8.5 Concluding Comments

In today’s society the police perform many functions and face many pressures. To reinforce their legitimacy with constituents governments rely on police to maintain order (Bayley, 1994b) and to demonstrate that the rule of law exists (Silberman, 1978). Citizens expect the police to keep them safe and to respond when called (Herbert, 2006). However, the politicised environment in which police operate can constrain the ability of police to respond to citizens in the most effective way (Buzawa, 2012; Felson et al., 2005).

Police seek to maintain and enhance their own organisational legitimacy, both with citizens and government, thereby securing status and resources (Herbert, 2006; Lawton et al., 2000; Tyler & Huo, 2002). However, demonstrating responsiveness to calls for service from citizens requires that police manage numerous challenges and pressures (Corboy et al., 2005; Gunal et al., 2007; Kuhn & Hoey, 1987; Moskos, 2007; Sacks et al., 1993; Sparrow et al., 1990). With so many varied expectations and challenges it is important that police utilise what mechanisms they can to manage this complexity.

Making effective use of performance measurement is one way police can improve their effectiveness by making sense of what is happening in the complex organisational context in which they operate (Gill, 2011; Speklé & Verbeeten, 2014). In this thesis I found that performance measurement is pervasive in police and is used to help reinforce police legitimacy with government and citizens (DiMaggio & Powell, 1983; Taylor, 2009; Willis et al., 2007). However, I also found that police appear to make only limited use of performance information to influence decisions that affect the quality of the service they deliver to citizens. Barriers exist to police managers making more effective use of performance information (Andersson & Tengblad, 2009; Coleman, 2008; Van Dooren et
al., 2010). However, opportunities also exist to stimulate more effective use of performance information by embedding analytical capability in police organisations, and increasing external scrutiny (Bayley, 2008; Bourdeaux & Chikoto, 2008; Dull, 2009; Skogan & Frydl, 2003).
References


Chan, J. B. L. (2001). The technological game: How information technology is transforming police practice Criminology and Criminal Justice, 1(2), 139-159.


Coleman, T. G. (2008). Managing strategic knowledge in policing: do police leaders have sufficient knowledge about organisational performance to make informed


http://www.homeoffice.gov.uk/rd s/recordedcrime1.html


Reeder, D. J. (1999). Impact of DNA typing on standards and practice in the forensic community. *Archives of Pathology and Laboratory Medicine, 123*(11), 1063-1065.


Appendix A: Informed Consent Package

The following four pages contain information provided to New Zealand Police interviewees prior to being interviewed. Similar information was provided to interviewees from Queensland Police Service. The differences between these versions are:

- references to New Zealand Police in the attached version were changed to ‘Queensland Police Service’ in the version given to Queensland Police Service personnel, and
- the New Zealand Police version contains a declaration regarding the fact that the researcher was an employee of New Zealand Police.

After reading the information supplied, interviewees were asked if they had any questions. They were then asked to sign the consent form on the last page of the information pack. After signing the consent form, the interview commenced.
INFORMATION SHEET

Who is conducting the research

Name: Gavin Knight

Research Centre and School: Centre of Excellence in Policing and Security
School of Criminology and Criminal Justice
Griffith University, Queensland, Australia

Contact Phone: Australia 0406 956646
New Zealand 027 2511176

Contact Email: g.knight@griffith.edu.au

Why is the research being conducted?

This research aims to better understand the nature of how police manage the performance of the service they provide in responding to calls for service from the public. It is being undertaken as part of a PhD research programme by the researcher named above.

What you will be asked to do

You will be asked to be interviewed in confidence about your knowledge of and views regarding the process of responding to calls for service from the public, and the associated performance measurement mechanisms used at various levels in the police agency you work for.

The basis by which participants will be selected or screened

The researcher has supplied your employer with descriptions of the types of people he needs to interview. These requirements include the type of role the person works in, the length of experience they have in this role, and that New Zealand Police (NZP) is willing for the person to interviewed by the researcher, should they consent to do so.

You have been identified, along with a number of others, as meeting the requirements for participation in this research. Other employees in a similar role are also being invited to participate.

The expected benefits of the research

This research aims to produce better understanding of how the performance of police response to calls for service from the public is monitored, and how this assists or otherwise influences the effectiveness of police in responding to calls for service from the public.
It seeks to identify aspects of these frameworks that managers and staff see as beneficial, and aspects they see as counterproductive. It also seeks to better understand what managers and staff consider constitutes effectiveness in responding to calls for service from the police. It is hoped that the understanding gained from this research will assist police in designing performance measurement frameworks that work well for staff and managers, supporting them to deliver an effective service to the public.

**Risks to you**

Risks to you are expected to be negligible. You will be asked questions by the researcher, to which you are free to answer as you wish. Your answers will later be analysed by the researcher, along with answers from other participants, in order to identify common themes. Results will be published in a way that does not identify individual participants.

You may choose whether your interview takes place in a private meeting room at your place of work, or offsite at a place like a private meeting room at a university campus.

**Your confidentiality**

Your interview will be recorded. After the interview, the recording will be transcribed by the researcher, then the recording will be destroyed.

In order to protect the anonymity of participants such as yourself, data will be coded, and the transcription and identifying code will be stored separately. In this way, the transcription of your interview will not contain identifying information such as your name, work-group, or date of interview.

Although de-identified quotes from interviews may be published, only the researcher, his two academic supervisors from Griffith University, and the external examiner of his PhD thesis, will be permitted access to raw transcripts of entire interviews. These transcripts will be stored securely until they are eventually destroyed, in line with the policies of Griffith University. NZP will not be granted access to these transcripts or recordings, and has agreed to this restriction.

The researcher is a NZP employee. However, NZP has agreed that this fact gives it no additional rights to information obtained by the researcher in this project. NZP has agreed to a confidentiality protocol that restricts it from compelling or pressuring the researcher to reveal not only what you say, but even whether or not you agreed to participate. Furthermore, the researcher will not share with anyone else in NZP, either your identity or information that could identify you.

**Your participation is voluntary**

Your participation is voluntary. Your decision whether or not to participate will in no way impact upon your relationship with your employer. Furthermore, you are free to withdraw from the study at any time.
Questions / further information

Please feel free to contact the researcher if you require additional information about the project.

Alternatively, if you require clarification from NZP, you may contact Superintendent Paula Rose, National Manager Road Policing. Superintendent Rose has been appointed by NZP as the liaison officer for this project. She can confirm both that this research has been approved by NZP, and that NZP has agreed to the protocols outlined to you in this document.

The ethical conduct of this research

Griffith University conducts research in accordance with Australia’s *National Statement on Ethical Conduct in Research Involving Humans*. If you have any concerns or complaints about the ethical conduct of the research project you should contact the Manager, Research Ethics on (++61 7) 3735 5585 or research-ethics@griffith.edu.au.

Feedback to you

It is intended that the research findings be published in a PhD thesis. Unless you indicate otherwise, the researcher will also write to you, summarising the key findings of the research.

Privacy Statement

The conduct of this research involves the collection, access and/or use of your identified personal information. The information collected is confidential and will not be disclosed to third parties without your consent, except to meet government, legal or other regulatory authority requirements. A de-identified copy of this data may be used for other research purposes. However, your anonymity will at all times be safeguarded. For further information consult the University’s Privacy Plan at www.griffith.edu.au/ua/aa/vc/pp or telephone (++61) (+7) 3735 5585.
The effectiveness of police in responding to calls for service

CONSENT FORM

Research Team

Name(s)  Gavin Knight

Research Centre  Centre of Excellence in Policing and Security
and School  School of Criminology and Criminal Justice
Griffith University, Queensland, Australia

Contact Phone  Australia 0406 956646
  New Zealand 027 2511176

Contact Email  g.knight@griffith.edu.au

By signing below, I confirm that I have read and understood the information package and in particular have noted that:

- I understand that my involvement in this research will include being interviewed by the researcher;
- I have had any questions answered to my satisfaction;
- I understand the risks involved;
- I understand that there will be no direct benefit to me from my participation in this research;
- I understand that my participation in this research is voluntary;
- I understand that if I have any additional questions I can contact the research team;
- I understand that I am free to withdraw at any time, without comment or penalty;
- I understand that I can contact the Manager, Research Ethics, at Griffith University Human Research Ethics Committee on 3735 5585 (or research-ethics@griffith.edu.au) if I have any concerns about the ethical conduct of the project; and
- I agree to participate in the project.

Name
Signature
Date
Appendix B: Interview Schedules

This appendix contains copies of the four interview schedules used to interview personnel from Queensland Police Service. Four almost identical schedules were used when interviewing personnel from New Zealand Police. In the schedules used to interview personnel from New Zealand Police, ‘Queensland Police Service’ was replaced with ‘New Zealand Police’.

As discussed in Chapter 5, the four schedules are as follows:

- **Frontline** – Used to interview Communications Room Operators and General Duties Officers
- **Middle Managers** – Used to interview PCC Managers and District Managers
- **Senior Executives** – Used to interview Senior Executives
- **Performance Managers** – Used to interview Performance Managers
Semi-structured Interview Questions: Frontline (Queensland)

[Note to interviewer: After gaining informed consent and advising the interviewee that the recording is commencing, for the recording, state the Interviewee’s Number, but not their name.]

1. **What were some of the jobs you did during the last few weeks that are typical of the sort of work you do in your current position?**

   [Identify only those specific jobs that are involved in police responding to a call for service from the public. If the interviewee identifies less than three such jobs, ask him/her to consider the previous week as well]

2. **Now I’d like to concentrate in particular on jobs that involve police response to calls for service from the public. This means all the things that police do from when a member of the public either requests help from police or reports to police a matter requiring police action, through to the end of police initial attendance at the scene, should police decide this is required. This excludes subsequent investigations, prosecution of alleged offenders, crime prevention initiatives, or other functions not related to police initial response to calls for service from the public.**

   So, I’d like you to think about the following jobs from the list that you just identified [list the relevant jobs from the response given to question 1.]

   [If the role that the interviewee has/took is unclear from what the interviewee has said so far, then ask the following question:]

   **Starting with [List one of the jobs], please tell me what your role was; how you received the job, and what actions you took.**

   [Repeat this question for each of the jobs identified by the interviewee that relate to police responding to calls for service from the public.]

3. **Again, thinking about each of these jobs one at a time, would you say the [List a job] job was successful?** [Ask this about each job separately.]

   [For each job, if the interviewee does not tell you why, ask:]

   3.1 Why was that (not) a successful job?
4. In Queensland Police Service, do you know if police measure their performance in responding to or managing these jobs?

[The interviewee may ask what you mean by “police”; Do you mean senior management, front-line police, or something else? Tell him/her – “Any of these”.

[If the interviewee does not know if police measure their performance, skip ahead to question 8; if he/she believes performance is measured, skip to question 6; if he/she believes performance is not measured, ask question 5.]

5. Do you know why police do not measure their performance in responding to or managing these jobs?

[If the interviewee answers yes, but doesn’t say why, prompt him/her as follows:]

5.1 Why do police not measure their performance in responding to or managing these jobs?

[Then skip to question 9]

6. Do you know why police measure their performance in responding to or managing these jobs?

[If the interviewee answers yes, but doesn’t say why, prompt him/her as follows:]

6.1 Why do police measure their performance in responding to or managing these jobs?

[This questionnaire does not include question 7, so progress to question 8.]

8. Do you know what the performance measures are that police use for these jobs or any aspect of their responding to calls for service?

[If the interviewee answers yes, but doesn’t say what they are, prompt him/her as follows:]

8.1 What are the performance measures?

9. As far as you are aware, who, either within or external to Queensland Police Service, receives reports on the performance of police in responding to calls for service from the public?

10. [Ask separately for each stakeholder identified in question 9]

Do you know what …………….. is looking to find out from these reports?
If the interviewee answers yes, but doesn’t elaborate sufficiently, prompt him/her as follows:

10.1 What is …………….. looking to find out from these reports?

11. What do you personally think matters in police responding effectively to calls for service from the public?

12. How do you know whether things are working well or not in police responding to calls for service from the public?

[This questionnaire does not include questions 13, and 14]

15. Still thinking about police response to calls for service from the public, can you think of a time when performance was assessed as being not good enough? This could either be in terms of police’s performance measures, or because of the reactions of stakeholders such as police managers, police staff, the public, or politicians.

If the interviewee answers ‘no’, skip on to question 17; if ‘yes’, ask:

16. What, if anything, happened as a result of performance being assessed as not good enough?

If the interviewee does not elaborate sufficiently, prompt him/her as required, by using questions such as:

16.1 What happened to the police organisation? or

16.2 What happened to staff or managers involved?

17. Are you involved in any way with reporting either your own performance, or the performance of any part of the process of police responding to calls for service?

If the interviewee answers ‘no’, move on to question 18; if ‘yes’, but he/she doesn’t go on to explain how, prompt him/her as follows:

17.1 What has been your involvement in reporting this performance?

17.2 What do you report about performance to management or others?

18. Earlier in this interview you told me about some jobs you have had in the last few weeks that relate to the police process of responding to calls for service. As
far as you know, are you involved in any way with recording information that is used by others to monitor performance about such jobs or process?

19. Is the performance monitoring that police do helpful?

[If the interviewee does not know, move on to question 20, otherwise, if he/she has not already told you why, ask the appropriate question as follows:]

19.1 How does this performance monitoring help? or
19.2 Why is this performance monitoring unhelpful?

[Then ask:]

19.3 Can you give me an example of how performance monitoring has either affected you or influenced what you do?

[This questionnaire does not include questions 20 and 21]

22. How does what you think of this performance monitoring influence what you record or report?

[If the interviewee’s answer to question 22 is that it has no influence, move to question 24; otherwise, unless he/she has provided a specific example, use the following question to elicit one:]

22.1 Can you give me an example of an instance where what you think about performance monitoring has influenced what you record or report?

[This questionnaire does not include question 23]

24. How do you know if your superiors think your performance is good?

[The interviewee, may say they don’t know, or my say something like “because my boss tells me when I’m not”. Probe deeper, to try to find out what understanding exists between the interviewee and their supervisor about what constitutes good performance for the supervisor. If no such understanding is apparent, seek to understand why and what impact this has.]

24.1 What does your supervisor monitor or measure, to assess whether you or your immediate team is performing well?
24.2 In the last year or two, what has your supervisor told you about what they want from you, in terms of performance?

25. **What do you do to make sure your supervisor is happy with your performance?**

26. **I have finished the set questions that I have for you. Do you have anything you wish to add that might be useful?**

   End.
Semi-structured Interview Questions: Middle Managers (Queensland)

[Note to interviewer: After gaining informed consent and advising the participant that the recording is commencing, for the recording, state the Participant’s Number, but not their name.]

Preamble

In this interview I’d like to concentrate in particular on police response to calls for service from the public. This means all the things that police do from when a member of the public either requests help from police or reports to police a matter requiring police action, through to the end of police initial attendance at the scene, should police decide this is required. This excludes subsequent investigations, prosecution of alleged offenders, crime prevention initiatives, or other functions not related to police initial response to calls for service from the public.

Do you understand what I mean, or would you like me to clarify any of this?

7. Do you have knowledge about any specific jobs involving police response to calls for service from the public, that have occurred in the last few weeks, and that involved either staff under your command or the area that you are responsible for?

[Prompt the participant to identify specific examples, rather than simply say “yes” or “no”.

3. Thinking about each of these jobs one at a time, would you say the [List a job] job was successful? [Ask this about each job separately.]

[For each job that the participant states they know whether or not the job was successful, if he/she does not tell you why, ask:]

3.1 Why was that (not) a successful job?

[The participant may answer this question in terms of either the outcome or in terms of what police did that caused a successful/unsuccesful outcome to result. Whichever perspective he/she gives, prompt him/her for the other perspective as well. I.e.]

3.2 What was successful/unsuccesful about the outcome? Or
3.3 What happened to cause the outcome of the job to succeed/fail?

4. **In QPS (Queensland Police Service), do you know if police measure their performance in responding to or managing these jobs?**

   [The participant may ask what you mean by “police”; Do you mean senior management, front-line police, or something else? Tell them—“Any of these”.

   [If the participant does not know if police measure their performance, skip ahead to question 8; if he/she believes performance is measured, skip to question 6; if he/she believes performance is not measured, ask question 5.]

5. **Do you know why police do not measure their performance in responding to or managing these jobs?**

   [If the participant answers yes, but doesn’t say why, prompt him/her as follows:]  
   
   5.1 Why do police not measure their performance in responding to or managing these jobs?

   [Then skip to question 9]

6. **Do you know why police measure their performance in responding to or managing these jobs?**

   [If the participant answers yes, but doesn’t say why, prompt him/her as follows:]  
   
   6.1 Why do police measure their performance in responding to or managing these jobs?

   [This questionnaire does not include question 7]

8. **Do you know what the performance measures are that police use for these jobs or any aspect of their responding to calls for service?**

   [If the participant answers yes, but doesn’t say what they are, prompt him/her as follows:]  
   
   8.1 What are the performance measures?

9. **As far as you are aware, who, either within or external to QPS, receives reports on the performance of police in responding to calls for service from the public?**
10. [Ask separately for each stakeholder identified in question 9]

Do you know what …………… is looking to find out from these reports?

[If the participant answers yes, but doesn’t elaborate sufficiently, prompt him/her as follows:]

10.1 What is …………… looking to find out from these reports?

10.2 How do you know this is what they are looking to find out from these reports?

11. What do you personally think matters in police responding effectively to calls for service from the public?

12. How do you know whether things are working well or not in police responding to calls for service from the public?

13. What do you find most helpful to monitor or measure?

14. What do you find most frustrating about what you can or cannot monitor or measure?

15. Still thinking about police response to calls for service from the public, can you think of a time when performance was assessed as being not good enough? This could either be in terms of police’s performance measures, or because of the reactions of stakeholders such as police managers, police staff, the public, or politicians.

[If the participant answers ‘no’, skip on to question 17; if ‘yes’, ask:]

16. What, if anything, happened as a result of performance being assessed as not good enough?

[If the participant does not elaborate sufficiently, prompt him/her as required, by using questions such as:]

16.1 What happened to the police organisation? or

16.2 What happened to staff or managers involved?

17. Are you or your staff involved in any way with reporting either the performance of your command (the area you are responsible for), or the performance of any part of the process of police responding to calls for service?
[If the participant answers ‘no’, move on to question 18; if ‘yes’, but he/she doesn’t go on to explain how, prompt him/her as follows]

17.1 What has been your involvement in reporting this performance?

17.2 What do you report about performance to senior management or others?

18. Earlier in this interview you told me about some jobs you have had in the last few weeks that relate to the police process of responding to calls for service. As far as you know, are you or your staff involved in any way with recording information that is used by others to monitor performance about such jobs or process?

19. Is the performance monitoring that police do helpful?

[If the participant does not know, move on to question 22, otherwise, if he/she has not already told you why, ask the appropriate question as follows:]

19.1 How does this performance monitoring help? or

19.2 Why is this performance monitoring unhelpful?

[Then ask:]

19.3 Can you give me an example of how performance monitoring has either affected you or influenced what you do?

[This questionnaire does not include questions 20 and 21.]

22. What motivates your decisions about which performance information to report or not report to senior management or external stakeholders such as government or the public?

[If the participant’s answer to question 22 is that he/she has no involvement in such decisions, move to question 23; otherwise, unless he/she has provided a specific example, use the following question to elicit one:]

22.1 Can you give me an example of an instance where what you think about performance monitoring has influenced what you report?

23. Do you yourself monitor anything different from or additional to the formal performance measurement framework, in order to assess whether or not your
command (the area you are responsible for) is effective in its role in police response to calls for service from the public?

[If the participant answers “no”, move on to question 24; otherwise, if he/she hasn’t answered this already, ask:]

23.1 What do you monitor to do this, and how do you go about it?

24. How do you know if your superiors think your performance is good?

[The participant, may say they don’t know, or my say something like “because my boss tells me when I’m not”. Probe deeper, to try to find out what understanding exists between the participant and their more senior managers about what constitutes good performance for the senior managers. If no such understanding is apparent, seek to understand why and what impact this has.]

24.1 What does your manager monitor or measure, to assess whether you or your immediate team is performing well?

24.2 In the last year or two, what has your manager told you about what they want from you, in terms of performance?

25. What do you do to make sure your manager is happy with your performance?

26. I have finished the set questions that I have for you. Do you have anything you wish to add that might be useful?

End.
Semi-structured Interview Questions: Senior Executives (Queensland)

[Note to interviewer: After gaining informed consent and advising the participant that the recording is commencing, for the recording, state the Participant’s Number, but not their name.]

Preamble

In this interview I’d like to concentrate in particular on police response to calls for service from the public. This means all the things that police do from when a member of the public either requests help from police or reports to police a matter requiring police action, through to the end of police initial attendance at the scene, should police decide this is required. This excludes subsequent investigations, prosecution of alleged offenders, crime prevention initiatives, or other functions not related to police initial response to calls for service from the public.

Do you understand what I mean, or would you like me to clarify any of this?

1. Do you have knowledge about any specific jobs involving police response to calls for service from the public, that have occurred in the last few weeks within Queensland Police Service?

[Prompt the participant to identify specific examples, rather than simply say “yes” or “no”. If the participant cannot think of at least three examples in the few weeks, ask about the previous six months.]

[This questionnaire does not include question 2]

3. Thinking about each of these jobs one at a time, would you say the [List a job] job was successful? [Ask this about each job separately.]

[For each job that the participant states they know whether or not the job was successful, if he/she does not tell you why, ask:]

3.1 Why was that (not) a successful job?

[The participant may answer this question in terms of either the outcome or in terms of what police did that caused a successful/unsucessful outcome to result. Whichever perspective he/she gives, prompt him/her for the other perspective as well. I.e.]
3.2 What was successful/unsuccessful about the outcome? Or
3.3 What happened to cause the outcome of the job to succeed/fail?

4. **In QPS (Queensland Police Service), do you know if police measure their performance in responding to or managing these jobs?**

[The participant may ask what you mean by “police”; Do you mean senior management, front-line police, or something else? Tell them – “Any of these”.]

[If the participant does not know if police measure their performance, skip ahead to question 8; if he/she believes performance is measured, skip to question 6; if he/she believes performance is not measured, ask question 5.]

5. **Do you know why police do not measure their performance in responding to or managing these jobs?**

[If the participant answers yes, but doesn’t say why, prompt him/her as follows:]

5.1 Why do police not measure their performance in responding to or managing these jobs?

[Then skip to question 9]

6. **Do you know why police measure their performance in responding to or managing these jobs?**

[If the participant answers yes, but doesn’t say why, prompt him/her as follows:]

6.1 Why do police measure their performance in responding to or managing these jobs?

[This questionnaire does not include question 7]

8. **Do you know what the performance measures are that police use for these jobs or any aspect of their responding to calls for service?**

[If the participant answers yes, but doesn’t say what they are, prompt him/her as follows:]

8.1 What are the performance measures?

9. **As far as you are aware, who, either within or external to QPS, receives reports on the performance of police in responding to calls for service from the public?**
10. [Ask separately for each stakeholder identified in question 9]

Do you know what ……………… is looking to find out from these reports?

[If the participant answers yes, but doesn’t elaborate sufficiently, prompt him/her as follows:]

10.1 What is ……………… looking to find out from these reports?

10.2 How do you know this is what they are looking to find out from these reports?

11. What do you personally think matters in police responding effectively to calls for service from the public?

12. How do you know whether things are working well or not in police responding to calls for service from the public?

13. What do you find most helpful to monitor or measure?

14. What do you find most frustrating about what you can or cannot monitor or measure?

15. Still thinking about police response to calls for service from the public, can you think of a time when performance was assessed as being not good enough? This could either be in terms of police’s performance measures, or because of the reactions of stakeholders such as police managers, police staff, the public, or politicians.

[If the participant answers ‘no’, skip on to question 19; if ‘yes’, ask:]  

16. What, if anything, happened as a result of performance being assessed as not good enough?

[If the participant does not elaborate sufficiently, prompt him/her as required, by using questions such as:]  

16.1 What happened to the police organisation? or

16.2 What happened to staff or managers involved?

[This questionnaire does not include questions 17 and 18]

19. Is the performance monitoring that police do helpful?
[If the participant does not know, move on to question 22, otherwise, if he/she has not already told you why, ask the appropriate question as follows:]

19.1 How does this performance monitoring help? or
19.2 Why is this performance monitoring unhelpful?

[Then ask:]

19.3 Can you give me an example of how performance monitoring has either affected you or influenced what you do?

[This questionnaire does not include questions 20 and 21]

22. What motivates your decisions about which performance information to report or not report to external stakeholders such as government or the public?

[If the participant’s answer to question 22 is that he/she has no involvement in such decisions, move to question 23; otherwise, unless he/she has provided a specific example, use the following question to elicit one:]

22.1 Can you give me an example of an instance where what you think about performance monitoring has influenced what you report?

23. Do you yourself monitor anything different from or additional to the formal performance measurement framework, in order to assess whether or not QPS is effective in the service it provides in responding to calls for service from the public?

[If the participant answers “no”, move on to question 24; otherwise, if he/she hasn’t answered this already, ask:]

23.1 What do you monitor to do this, and how do you go about it?

24. How do you know if the minister of police and other key stakeholders think your performance is good?

24.1 What do the minister of police and other key stakeholders to whom you are accountable, monitor or measure, to assess whether you are performing well?

24.2 In the last year or two, what has the minister of police or other key stakeholders to whom you are accountable, told you about what they want from you, in terms of performance in the area of police response to calls for service?
25. What do you do to make sure the minister of police and other key stakeholders to whom you are accountable, are happy with your performance?

26. I have finished the set questions that I have for you. Do you have anything you wish to add that might be useful?

End.
Semi-structured Interview Questions: Performance Managers (Queensland)

[Note to interviewer: After gaining informed consent and advising the participant that the recording is commencing, for the recording, state the Participant’s Number, but not their name.]

Preamble

In this interview I’d like to concentrate in particular on police response to calls for service from the public. This means all the things that police do from when a member of the public either requests help from police or reports to police a matter requiring police action, through to the end of police initial attendance at the scene, should police decide this is required. This excludes subsequent investigations, prosecution of alleged offenders, crime prevention initiatives, or other functions not related to police initial response to calls for service from the public.

Do you understand what I mean, or would you like me to clarify any of this?

[This questionnaire does not include questions 1, 2, and 3]

4. Does the Queensland Police Service measure its performance in its response to calls for service from the public?

[Note: The participant may ask about what level or group I am referring to when I say “Queensland Police Service”. i.e. Who do I mean? Answer that this can be any group or at any level you can think of.]

[If the participant does not know if police measure their performance, skip ahead to question 8; if he/she believes performance is measured, skip to question 6; if he/she believes performance is not measured, ask question 5.]

5. Why does QPS not measure its performance in responding to calls for service from the public? [then skip to question 9]

6. Why does QPS measure its performance in responding to calls for service from the public?

7. How does QPS measure its performance in responding to calls for service from the public?
8. Do you know what the performance measures are that QPS uses in relation to police response to calls for service from the public?

[If the participant answers yes, but doesn’t say what they are, prompt him/her as follows:]

8.1 What are these performance measures?

9. As far as you are aware, who, either within or external to QPS, receives reports on the performance of police in responding to calls for service from the public?

10. [Ask separately for each stakeholder identified in question 9]

   Do you know what ................ is looking to find out from these reports?

[If the participant answers yes, but doesn’t elaborate sufficiently, prompt him/her as follows:]

10.1 What is ................ looking to find out from these reports?

10.2 How do you know this is what they are looking to find out from these reports?

11. What do you personally think matters in police responding effectively to calls for service from the public?

12. How do you know whether things are working well or not in police responding to calls for service from the public?

13. What do you find most helpful to monitor or measure?

14. What do you find most frustrating about what you can or cannot monitor or measure?

15. Still thinking about police response to calls for service from the public, can you think of a time when performance was assessed as being not good enough? This could either be in terms of police’s performance measures, or because of the reactions of stakeholders such as police managers, police staff, the public, or politicians.

[If the participant answers ‘no’, skip to question 19; if ‘yes’, ask question 16.]

16. What, if anything, happened as a result of performance being assessed as not good enough?
[If the participant does not elaborate sufficiently, prompt him/her as required, by using questions such as:]

116.1 What happened to the police organisation?

216.2 What happened to staff or managers involved? or

316.3 What happened to the performance management framework?

[This questionnaire does not include questions 17 and 18]

19. Is the performance monitoring that police do helpful?

[If the participant does not know, move on to question 20, otherwise, if he/she has not already told you why, prompt with the appropriate question as follows:]

19.1 How does this performance monitoring help? or

19.2 Why is this performance monitoring unhelpful?

[Then ask:]

19.3 Can you give me an example of how performance monitoring has either affected you or influenced what you do?

20. What problems have you encountered in monitoring police performance in responding to calls for service from the public?

21. Are you aware of changes that have been made either to performance measures or to the performance measurement framework, because of previous problems in measuring or reporting performance.

[If the participant answers yes, but doesn’t say what they are, prompt him/her as follows:]

21.1 What were these problems, and what changes were made as a result?

[This questionnaire does not include questions 22, 23, 24, and 25]

26. I have finished the set questions that I have for you. Do you have anything you wish to add that might be useful?

End.
Appendix C: Results of Tests Performed in Study 2

This appendix presents all Granger-causality test results produced in step 5 of Study 2 (see section 6.7). This is possible because Granger-causality test results can be presented in a compact form. However, cross-correlograms produced for step 4 of Study 2 (see section 6.6) are charts that require more space. Therefore, only a few illustrative examples of cross-correlograms are presented herein. Specifically, the nine cross-correlograms corresponding to the nine tests performed on New Zealand data for the relationship between demand volumes (DV) and general response-time performance (GR) are presented. These nine cross-correlograms include one for each of the three types of model for each of the three New Zealand Police districts, and constitute part of the analysis undertaken to test Hypothesis 1.

The cross-correlogram corresponding for the test performed on Queensland ARIMA modelled data for the relationship between sick leave (SL) and general response-time performance (GR) in District 6 is also presented. This cross-correlogram is the only one of the 45 used to test Hypothesis 4 that shows a significant cross-correlation at lag 0. As discussed in section 7.4.4, unlike Hypotheses 1, 2, and 3, Hypothesis 4 requires testing relationships at lag 0.

**C1: Key**

The correlograms presented in this appendix are screenshots from EViews, so show the variables names used in EViews. Names of variables used in EViews differ from those used in Chapter 7, as they differentiate between time series pre-whitened using different types of model. The first three characters of the EViews variable name identifies the variable referred to in Chapter 7. Table C.1 provides a translation between the first three
characters of each EViews variable name and name given to the same variable in Chapter 7.

Table C.1
Translation Between Names of Variables in EViews and Chapter 7

<table>
<thead>
<tr>
<th>EViews variable name</th>
<th>Chapter 7 variable name</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC</td>
<td>CV</td>
</tr>
<tr>
<td>DEE</td>
<td>DV</td>
</tr>
<tr>
<td>DRH</td>
<td>ER</td>
</tr>
<tr>
<td>DRM</td>
<td>GR</td>
</tr>
<tr>
<td>DHH</td>
<td>RC</td>
</tr>
<tr>
<td>HHH</td>
<td>SH</td>
</tr>
<tr>
<td>HSS</td>
<td>SL</td>
</tr>
<tr>
<td>OOU</td>
<td>OV</td>
</tr>
<tr>
<td>ORUR</td>
<td>RR</td>
</tr>
<tr>
<td>PTI</td>
<td>NV</td>
</tr>
</tbody>
</table>

All variables referred to in this appendix are pre-whitened time series produced by step 3 of the transfer function analysis methodology described in section 6.5. EViews variable names include a suffix that identifies which type of model was used to pre-whiten the series. Time series pre-whitened using an ARIMA model have the suffix RES; time series pre-whitened using a VAR-AIC model have the suffix _W; time series pre-whitened using a VAR-12 model have the suffix _W12. Using this convention, the time series for DV pre-whitened using an ARIMA model is named DEERES in EViews cross-correlograms presented in section C.2. Similarly, the time series for GR pre-whitened using a VAR-AIC model is named DRM_W in the EViews cross-correlograms.
C2: Illustrative examples of cross-correlation functions

Figure C.1
Cross-correlograms for DV-GR in District 1

<table>
<thead>
<tr>
<th>District 1</th>
<th>DV-GR (DEE-DRM)</th>
<th>Granger tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEERES,DRMRES(-i)</td>
<td>DEERES,DRMRES(+i)</td>
<td>Not performed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>lag</td>
<td>lead</td>
</tr>
<tr>
<td>0</td>
<td>-0.0178</td>
<td>-0.0178</td>
</tr>
<tr>
<td>1</td>
<td>-0.3357</td>
<td>0.0434</td>
</tr>
<tr>
<td>2</td>
<td>-0.2552</td>
<td>-0.0429</td>
</tr>
<tr>
<td>3</td>
<td>0.1287</td>
<td>0.2017</td>
</tr>
<tr>
<td>4</td>
<td>0.0365</td>
<td>0.0767</td>
</tr>
<tr>
<td>5</td>
<td>-0.0660</td>
<td>0.0590</td>
</tr>
<tr>
<td>6</td>
<td>0.0150</td>
<td>-0.1431</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEE_W,DRM_W(-i)</th>
<th>DEE_W,DRM_W(+i)</th>
<th>Not performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>lag</td>
<td>lead</td>
</tr>
<tr>
<td>0</td>
<td>0.0885</td>
<td>0.0885</td>
</tr>
<tr>
<td>1</td>
<td>-0.3190</td>
<td>-0.1324</td>
</tr>
<tr>
<td>2</td>
<td>-0.1895</td>
<td>-0.0775</td>
</tr>
<tr>
<td>3</td>
<td>0.1838</td>
<td>0.0139</td>
</tr>
<tr>
<td>4</td>
<td>-0.0384</td>
<td>0.0919</td>
</tr>
<tr>
<td>5</td>
<td>-0.0326</td>
<td>0.0895</td>
</tr>
<tr>
<td>6</td>
<td>0.0080</td>
<td>-0.0844</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEE_W12,DRM_W12(-i)</th>
<th>DEE_W12,DRM_W12(+i)</th>
<th>Not performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>lag</td>
<td>lead</td>
</tr>
<tr>
<td>0</td>
<td>0.0538</td>
<td>0.0538</td>
</tr>
<tr>
<td>1</td>
<td>-0.2026</td>
<td>-0.0180</td>
</tr>
<tr>
<td>2</td>
<td>-0.1209</td>
<td>-0.1397</td>
</tr>
<tr>
<td>3</td>
<td>0.0948</td>
<td>0.1141</td>
</tr>
<tr>
<td>4</td>
<td>-0.0020</td>
<td>0.1228</td>
</tr>
<tr>
<td>5</td>
<td>-0.1267</td>
<td>0.1671</td>
</tr>
<tr>
<td>6</td>
<td>-0.0625</td>
<td>0.0432</td>
</tr>
</tbody>
</table>

The vertical dotted line in the cross-correlogram represents two standard error bounds. The right-hand-side (lead) of the cross-correlogram represents the correlation between DV (DEE) and a future value of GR (DRM). The above cross-correlograms for District 1 show that no cross-correlation coefficient for any lag (DV leading GR) exceeds two standard error bounds. No Granger-causality tests were performed for these tests because no correlation coefficients in the cross-correlogram exceeded two standard error bounds.
The above cross-correlograms for District 2 show that only the ARIMA model cross-correlation coefficient for lag 1 (DV leading GR) exceeds two standard error bounds. The Granger-causality test for this relationship is significant at \( p = 0.05 \).
Figure C.3
Cross-correlograms for DV-GR in District 3

<table>
<thead>
<tr>
<th>District 3</th>
<th>DV-GR (DRM-DEE)</th>
<th>Granger tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>p &gt; 0.1</td>
</tr>
</tbody>
</table>

### Granger tests for District 3

<table>
<thead>
<tr>
<th>DEERES3,DRMRES3(-i)</th>
<th>DEERES3,DRMRES3(+i)</th>
<th>i</th>
<th>lag</th>
<th>lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0006</td>
<td>1</td>
<td>-0.2600</td>
<td>-0.1359</td>
</tr>
<tr>
<td>1</td>
<td>0.2316</td>
<td>2</td>
<td>0.0524</td>
<td>0.0033</td>
</tr>
<tr>
<td>3</td>
<td>0.0102</td>
<td>4</td>
<td>-0.0225</td>
<td>0.0599</td>
</tr>
<tr>
<td>5</td>
<td>0.1080</td>
<td>6</td>
<td>0.2746</td>
<td>0.0110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEE_W,DRM_W(-i)</th>
<th>DEE_W,DRM_W(+i)</th>
<th>i</th>
<th>lag</th>
<th>lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0786</td>
<td>1</td>
<td>-0.2247</td>
<td>-0.2156</td>
</tr>
<tr>
<td>2</td>
<td>0.2935</td>
<td>3</td>
<td>0.0431</td>
<td>-0.0010</td>
</tr>
<tr>
<td>4</td>
<td>-0.0614</td>
<td>5</td>
<td>0.0180</td>
<td>0.2023</td>
</tr>
<tr>
<td>6</td>
<td>-0.0795</td>
<td>6</td>
<td>0.0139</td>
<td>-0.1439</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEE_W12,DRM_W12(-i)</th>
<th>DEE_W12,DRM_W12(+i)</th>
<th>i</th>
<th>lag</th>
<th>lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.0495</td>
<td>1</td>
<td>-0.2761</td>
<td>-0.1602</td>
</tr>
<tr>
<td>2</td>
<td>-0.2169</td>
<td>3</td>
<td>-0.0308</td>
<td>0.0586</td>
</tr>
<tr>
<td>4</td>
<td>-0.1737</td>
<td>5</td>
<td>0.0717</td>
<td>0.1520</td>
</tr>
<tr>
<td>6</td>
<td>0.0140</td>
<td>6</td>
<td>-0.0152</td>
<td></td>
</tr>
</tbody>
</table>

Not performed

The above cross-correlograms for District 3 show that both the ARIMA and VAR-AIC model cross-correlation coefficients for lag 5 (DV leading GR) exceed two standard error bounds. The Granger-causality test for this relationship in the VAR-AIC model is significant at p = 0.05. However, the Granger-causality test for this relationship in the ARIMA model is not significant, even at p = 0.1.
The above cross-correlogram for SL-GR in District 6 illustrates that the correlation coefficient of 0.2986 at lag 0 exceeds two standard errors, represented by the dotted vertical lines. This is the only cross-correlogram of the 45 produced to test Hypothesis 4 that indicates a correlation coefficient at lag 0 that exceed two standard error bounds. Also, evident in this cross-correlogram is a marginally significant correlation coefficient at lag 1. Granger-causality test results for this and other pairs of variables are presented in section C.3.

**C3: Granger-causality test results**

I performed a Granger-causality test in all instances where a cross-correlation function in step 4 showed a correlation coefficient exceeding two standard error bounds at a lag between 1 and 6 months. Tables C.2 to C.25 present results for all Granger causality tests that rejected the null hypothesis with \( p < 0.1 \). Results are presented in the form ‘x/y’, where \( x \) indicates the lag and \( y \) indicates the p-value of the Granger-causality test. A minus sign preceding the lag number indicates the relationship is inverse. Lack of a minus sign indicates the relationship is reinforcing. In instances where more than one significant lag between 1 and 6 exists, results are presented for the first lag (least delay) where \( p < 0.5 \).
Results for tests of Hypothesis 1

Table C.2
Granger-causality Test Results for Hypothesis 1 – District 1

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-4/0.0265</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-</td>
<td>-1/0.0246</td>
</tr>
<tr>
<td>ER-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-4/0.0911</td>
<td>-4/0.0455</td>
</tr>
<tr>
<td>DV-SH</td>
<td>-</td>
<td>1/0.0475</td>
</tr>
<tr>
<td>OV-SH</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-SH</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RR</td>
<td>-2/0.0138</td>
<td>5/0.0295</td>
</tr>
</tbody>
</table>
Table C.3
Granger-causality Test Results for Hypothesis 1 – District 2

<table>
<thead>
<tr>
<th></th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-GR</td>
<td>-1/0.0477</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>DV-RC</td>
<td>-</td>
<td>2/0.0186</td>
<td>2/0.0092</td>
<td>Significant lag 2 reinforcing</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-RR</td>
<td>1/0.0103</td>
<td>1/0.0022</td>
<td>-</td>
<td>Significant lag 1 reinforcing</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-1/0.0103</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>ER-RC</td>
<td>2/0.0123</td>
<td>2/0.0173</td>
<td>2/0.0233</td>
<td>Significant lag 2 reinforcing</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>DV-SH</td>
<td>-5/0.0303</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-SH</td>
<td>6/0.0433</td>
<td>-</td>
<td>2/0.0285</td>
<td>Fail</td>
</tr>
<tr>
<td>ER-SH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SH-GR</td>
<td>-</td>
<td>-1/0.0581</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SH-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SH-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>
### Table C.4
Granger-causality Test Results for Hypothesis 1 – District 3

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>5/0.0832</td>
</tr>
<tr>
<td>DV-RC</td>
<td>-</td>
<td>-1/0.00347</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-4/0.0024</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-1/0.0383</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-1/0.0279</td>
<td>-1/0.0566</td>
</tr>
<tr>
<td>ER-RC</td>
<td>-6/0.0346</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-SH</td>
<td>-5/0.0660</td>
<td>-5/0.0272</td>
</tr>
<tr>
<td>ER-SH</td>
<td>-1/0.0110</td>
<td>-1/0.0007</td>
</tr>
<tr>
<td>SH-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SH-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table C.5
Granger-causality Test Results for Hypothesis 1 – District 4

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-2/0.0630</td>
<td>-2/0.0014</td>
</tr>
<tr>
<td>OV-GR</td>
<td>3/0.0468</td>
<td>3/0.0973</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-</td>
<td>-2/0.0974</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-1/0.0210</td>
<td>-1/0.0233</td>
</tr>
</tbody>
</table>
Table C.6
Granger-causality Test Results for Hypothesis 1 – District 5

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table C.7
Granger-causality Test Results for Hypothesis 1 – District 6

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-GR</td>
<td>-</td>
<td>-2/0.0326</td>
</tr>
<tr>
<td>DV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OV-RR</td>
<td>1/0.0111</td>
<td>1/0.0215</td>
</tr>
<tr>
<td>ER-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ER-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Results for tests of Hypothesis 2

Table C.8
Granger-causality Test Results for Hypothesis 2 – District 1

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-CV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>DV-NV</td>
<td></td>
<td>-</td>
<td>-1/0.0518</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-CV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-NV</td>
<td></td>
<td>-</td>
<td>-1/0.0401</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>ER-CV</td>
<td></td>
<td>-</td>
<td>-1/0.0692</td>
<td>-1/0.0625</td>
<td>Indeterminate lag 1 inverse</td>
</tr>
<tr>
<td>ER-NV</td>
<td></td>
<td>-2/0.0432</td>
<td>-2/0.0534</td>
<td>-</td>
<td>Indeterminate lag 2 inverse</td>
</tr>
</tbody>
</table>

Table C.9
Granger-causality Test Results for Hypothesis 2 – District 2

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV-CV</td>
<td></td>
<td>5/0.0439</td>
<td>-4/0.0426</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>DV-NV</td>
<td></td>
<td>-</td>
<td>-6/0.0523</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-CV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>OV-NV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>ER-CV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>ER-NV</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>
Table C.10
Granger-causality Test Results for Hypothesis 2 – District 3

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>DV-NV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>OV-NV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>ER-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>ER-NV</td>
<td>-6/0.0027</td>
<td>-</td>
</tr>
</tbody>
</table>

Table C.11
Granger-causality Test Results for Hypothesis 2 – District 4

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>ER-CV</td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

Table C.12
Granger-causality Test Results for Hypothesis 2 – District 5

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-CV</td>
<td>-6/0.0497</td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>ER-CV</td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>
Table C.13
Granger-causality Test Results for Hypothesis 2 – District 6

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>DV-CV</td>
<td>4/0.0105</td>
<td>-</td>
</tr>
<tr>
<td>OV-CV</td>
<td>-</td>
<td>-4/0.0893</td>
</tr>
<tr>
<td>ER-CV</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Results for tests of Hypothesis 3

Table C.14
Granger-causality Test Results for Hypothesis 3 – District 1

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>CV-GR</td>
<td>2/0.0978</td>
<td>2/0.0704</td>
</tr>
<tr>
<td>CV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-GR</td>
<td>-4/0.0164</td>
<td>-4/0.0331</td>
</tr>
<tr>
<td>NV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-RR</td>
<td>-</td>
<td>4/0.0342</td>
</tr>
</tbody>
</table>

Table C.15
Granger-causality Test Results for Hypothesis 3 – District 2

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA</td>
<td>VAR-AIC</td>
</tr>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-GR</td>
<td>-4/0.0002</td>
<td>-4/0.0000</td>
</tr>
<tr>
<td>NV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### Table C.16
Granger-causality Test Results for Hypothesis 3 – District 3

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-GR</td>
<td>-4/0.0002</td>
<td>-4/0.0010</td>
</tr>
<tr>
<td>NV-RC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NV-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table C.17
Granger-causality Test Results for Hypothesis 3 – District 4

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table C.18
Granger-causality Test Results for Hypothesis 3 – District 5

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>VAR-AIC</td>
<td>VAR-12</td>
</tr>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>2/0.0749</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table C.19
Granger-causality Test Results for Hypothesis 3 – District 6

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>C-V</th>
<th>GR</th>
<th>AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>CV-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>

Results for tests of Hypothesis 4

Table C.20
Granger-causality Test Results for Hypothesis 4 – District 1

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>C-V</th>
<th>GR</th>
<th>AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>5/0.0920</td>
<td>-3/0.0343</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SL-RC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>

Table C.21
Granger-causality Test Results for Hypothesis 4 – District 2

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>C-V</th>
<th>GR</th>
<th>AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>1/0.0142</td>
<td>-3/0.0309</td>
<td>-3/0.0473</td>
<td>-</td>
<td>-</td>
<td>Significant lag 3 inverse</td>
</tr>
<tr>
<td>SL-RC</td>
<td>-</td>
<td>4/0.0130</td>
<td>4/0.0485</td>
<td>-</td>
<td>-</td>
<td>Significant lag 4 reinforcing</td>
</tr>
<tr>
<td>SL-RR</td>
<td>3/0.0034</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>
Table C.22
Granger-causality Test Results for Hypothesis 4 – District 3

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>-</td>
<td>-2/0.0773</td>
<td>-2/0.0482</td>
<td>-</td>
<td>Indeterminate lag 2 inverse</td>
</tr>
<tr>
<td>SL-RC</td>
<td>-1/0.0267</td>
<td>-1/0.0217</td>
<td>-</td>
<td></td>
<td>Significant lag 1 inverse</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-4/0.0747</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>

Table C.23
Granger-causality Test Results for Hypothesis 4 – District 4

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>

Table C.24
Granger-causality Test Results for Hypothesis 4 – District 5

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>2/0.0365</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>

Table C.25
Granger-causality Test Results for Hypothesis 4 – District 6

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Model</th>
<th>ARIMA</th>
<th>VAR-AIC</th>
<th>VAR-12</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-GR</td>
<td>1/0.0352</td>
<td>-5/0.0239</td>
<td>-5/0.0509</td>
<td></td>
<td>Indeterminate lag 5 inverse</td>
</tr>
<tr>
<td>SL-RR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Fail</td>
</tr>
</tbody>
</table>