

CRANFIELD UNIVERSITY

KHALID A. AL-AKWA

**Adoption and Implementation of Performance Measurement System in the
Yemeni Public Sector: Challenges and Prospects**

SCHOOL OF MANAGEMENT

DOCTOR OF BUSINESS ADMINISTRATION THESIS

SUPERVISOR: MICHAEL BOURNE

JUNE 2017

CRANFIELD UNIVERSITY SCHOOL OF MANAGEMENT

Doctor of Business Administration 2017

KHALID A. AL-AKWA

**Adoption and Implementation of Performance Measurement System in the
Yemeni Public Sector: Challenges and Prospects**

Supervisors: MICHAEL BOURNE

June 2017

This thesis is submitted in partial fulfillment of the requirements for the degree of
Doctor of Business Administration.

© Cranfield University 2017. All rights reserved. No part of this publication may be
reproduced without the written permission of the copyright owner.

ABSTRACT

This thesis is submitted for the Cranfield DBA program. It examines and explores the utilization of Performance Measurement Systems (PMS) in public organizations with a special focus on the culture of developing countries. The research was inspired by the apparent potential of measurement for reform and the limited research on the implementation of performance measurement systems in the public sector. The paper started by conducting a thorough and critical examination of a vast body of literature using a systematic approach in order to identify the key determinants of successful utilization of PMS in public organizations. The review revealed that utilization of performance measurement in the public sector is a work in progress. It showed that, in spite of the growing interest on the field of performance measurement to improve public services, the empirical work remains small and disintegrated, focused only on single issues, and lacks a holistic perspective. The literature indicates that little is known about the needed conditions for an effective utilization of measurement within the public context in general and more so in non-western countries. The thesis, then, report the result of an empirical research based on in-depth interviews and focus group meetings with senior public managers from multiple public organizations in the Republic of Yemen. The empirical project sought to examine the Yemeni public sector organizations' familiarity with performance measurement and explore the actual and potential role of measurement to support reform and improve performance in the public sector. Findings indicate that consecutive reforms in the Yemeni public sector did not seem to emphasize the need to establish a performance-based culture. It shows that the Yemeni public sector features a traditional administration that is away from a result-based management, bound to roles and process, and less flexible to change.

This study, accordingly, highlights the significant importance of the context and subsequent processes and arrangements that promote or impede the development and use of performance measurement in public organizations in non-western countries. It shows that a one-size-fits-all policy should be avoided by policy makers who should be mindful of possible dissimilarities between public entities and the significance and relevance of the context. The study indicates that failure and success of utilization of PMS in the public sector is influenced by various elements that need to be considered prior and after the adoption of measurement. A new theoretical model has been developed based on findings in order to extend the existing literature in this area. The model provides an extensive description about the main factors affecting adoption and implementation of PMS and promotes understanding about the needed conditions for successful utilization in the public sector. It addressed the need for translating abstract concepts such as adoption and implementation into clearly defined theoretical constructs in order to avoid confusion and create a common vocabulary for researchers.

ACKNOWLEDGEMENTS

This study may not have been completed without the support and guidance of the following people. I am grateful to my supervisor, panel members, team of Cranfield University, colleagues, friends, and family, for their kind support, encouragement, and patience.

My sincere appreciation goes to my supervisor, Mike Bourne, for providing grateful support and guidance for the past years. Many people go through tough times but very few are lucky to have someone who understands and provide support when needed. For that, I will always be indebted.

Many thanks to all my panel members for their kind support and feedback. Special thanks to Dr. Jutta Tobias, for her sincere compassion and for offering excellent insights and constructive critique.

I am grateful to everyone in the School of Management at Cranfield University. I appreciate their responsiveness and kind assistance over the past years. To Barbara Waites, Alison Wilkerson, Carlos Mena, Alison Southgate, Emma Parry, and many others, thank you for your dedication and support. All students at the University are lucky to have people like you looking after them.

To all my classmates, I will never forget the wonderful times we had during our cohort weeks in Cranfield. Your success has been an important incentive for me to catch up.

To my wonderful friends, life would be so much different without you. Thanks for your advice and encouragement.

My deepest gratitude and sincere thanks goes to my family. To my father, I will never forget your tears of joy once you heard that I have finished my doctorate. You have been my inspiration. I will always do my best to make you proud. To my mother, your prayers and words of encouragement have been the light that shined my way and kept me going. To my loving wife, this is yours as it is mine. Thanks for your patience and support. I could have not possibly done it without you. To my three daughters: Abeer, Reem, and Noor, and to my brothers and sisters, thanks for the unconditional love and support, I am really blessed to have you in my life.

Finally, I thank all the anonymous people who supported and assisted me in completing this work. Your names might not be listed, but I will always remember what you did for me. I thank you for your time and guidance.

TABLE OF CONTENTS

ABSTRACT	3
ACKNOWLEDGEMENTS.....	5
TABLE OF CONTENTS	6
LIST OF TABLES AND FIGURES	10
CHAPTER 1: BACKGROUND AND RATIONALE FOR THE RESEARCH	11
Preface.....	12
Overview of the Research Process.....	14
Introduction	14
The Critical Review	15
The Qualitative Research	15
The Scoping Study	16
The Systematic Review	19
The Empirical Project.....	24
CHAPTER 2: A LITERATURE REVIEW OF PERFORMANCE MEASUREMENT SYSTEM IMPLEMENTATION AND USE IN PUBLIC ORGANIZATIONS	26
Introduction.....	27
General Overview	28
Change Management.....	28
Implementation of PMS	31
Public Management.....	34
Managing Change in Public Organizations.....	36
Performance Measurement	40
Measuring Performance in the Public Sector.....	42
Discussion.....	45
Systematic Review Questions.....	47
Scoping Study Summary Paper	48
CHAPTER 3: RESEARCH METHODOLOGY	52

Research Philosophy	53
Research Strategy and Design	55
Data Collection	57
Data Analysis	58
Ethical Considerations	59
Chapter Summary	60
CHAPTER 4: PERFORMANCE MEASUREMENT UTILIZATION IN THE PUBLIC SECTOR: A SYSTEMATIC REVIEW	62
1. Overview	63
2. Methodology	65
2.1 Introduction	65
2.2 Objectives.....	65
2.3 Systematic Review Back Ground.....	66
2.4 Process of the Systematic Review.....	66
2.5 Planning the Review.....	67
2.6 The Search Process.....	75
2.7 Final Result	78
2.8 Data Analysis and Synthesis	79
2.9 Descriptive and Thematic Analysis.....	80
3. Determinants of Successful Utilization of PMS	90
3.1 Introduction	90
3.2 Adopting PMS in the Public Sector	92
3.3 Development and Use of Measures- PI Supply and Consumption	96
3.4 Summary and Conclusions.....	132
Appendix.....	136
CHAPTER 5: THE PATTERNS OF PERFORMANCE MEASUREMENT APPLICATION IN THE YEMENI PUBLIC SECTOR	169
Introduction	170
Literature Review	172
Research Methods.....	178
<i>Introduction</i>	178
<i>Purposeful Sampling versus Randomness</i>	178
<i>Selecting the Sample</i>	180
<i>Data Collection Stages</i>	180

<i>Conducting the interviews</i>	182
<i>Data Analysis</i>	182
<i>Translation Issues</i>	184
The Yemeni Context.....	185
<i>Overview</i>	185
<i>Political and Economic Challenges</i>	185
<i>Reform in the Public Sector</i>	187
Findings and Analysis	189
<i>Understanding the concept</i>	189
<i>Source and Types of PI Demand and Supply</i>	192
<i>Main Barriers</i>	195
Discussion.....	203
Conclusions	210
Limitations and Further research.....	212
APPENDIX 1 – INTERVIEW SCHEDULE	214
CHAPTER 6: CONCEPTUAL FOUNDATION FOR A THEORY OF PMS UTILIZATION IN THE PUBLIC SECTOR	218
Overview.....	219
Theoretical Background.....	219
General Overview of the Conceptual Model	222
Elements of the Conceptual Model.....	226
The Utilization Issue	226
Factors Affecting Utilization of PMS	229
<i>Political factors</i>	229
Organizational Capacity	234
<i>Organizational Culture</i>	236
Interrelation between Processes and Factors	238
<i>Interrelation between processes</i>	238
<i>Interrelation between factors</i>	239
Summary and Conclusions	242
CHAPTER 7: KEY FINDINGS, CONTRIBUTIONS AND LIMITATIONS	244
Key Findings.....	245
Key Findings of Project 1.....	245
Key Findings of Empirical Project.....	249

Contributions.....	253
<i>Contribution to Knowledge</i>	253
<i>Contribution to Practice</i>	259
Limitations and Suggestions for Further Research.....	263
REFERENCES	267

LIST OF TABLES AND FIGURES

Table 1 Final Search String	19
Table 2 Empirical Findings	25
Table 3 Review Panel	68
Table 4 Initial Journal Selection	69
Table 5 The Search Databases	70
Table 6 Initial Search Strings	71
Table 7 Final Search String	72
Table 8 Papers' Relevance on a Four Points Scale	73
Table 9 Inclusion/ Exclusion Criteria	74
Table 10 Example of a search using papers' keywords	76
Table 11 Search Results	78
Table 12 Removing Redundant Papers	78
Table 13 SR sources	79
Table 14 Journals and ranking	81
Table 15 Development and Use of measures	99
Table 16 Public organizations by their types	180
Table 17 No. Interviewees by their title and type of organization	181
Table 18 Theoretical and Empirical Findings	225
Figure 1 Mapping the Fields	27
Figure 2 Cranfield Systematic Review Process	67
Figure 3 The Search Process	76
Figure 4 Literature Source	81
Figure 5 Papers by journal ranking	83
Figure 6 PMS design	84
Figure 7 Papers by the field of interest	84
Figure 8 Research strategy	85
Figure 9 Research strategy timeline	86
Figure 10 Papers by publication year	86
Figure 11 Performance measurement vs. performance management	87
Figure 12 Papers by the type of knowledge	88
Figure 13 Sources by Country	88
Figure 14 Demand of PI in public organizations	93
Figure 15 PI Supply and Consumption	97
Figure 16 PI demand, supply and use	134
Figure 17 Adapted from Julnes and Holzer Framework 2001	221
Figure 18 Adapted From Carlucci, Schiuma and Sole, 2015 Model	221
Figure 19 the Conceptual Model	224

CHAPTER 1: BACKGROUND AND RATIONALE FOR THE RESEARCH

Preface

Performance measurement has become a key component of modern governance in many countries around the world. Performance measurement systems (PMS) has been viewed as a dynamic tool that support and facilitates decision making process through the collection, analysis and reflection on individual and organizational performance data. As a result, many governments have adopted PMS in order to enhance efficiency and effectiveness of the public sector. However, the implementation of PMS in the public sector has faced various difficulties and problems that are not necessarily similar for each organization or country. The literature points to some of these problems such as, for instance, the proliferation of measures, overreliance on financial measures at the expense of other type of measures, measures do not reflect reality, focus on some easy-to-measure rather than hard-to-measure aspects, and measures developed for legitimacy purposes only and not for improving performance (Carlucci, Schiuma and Sole, 2015; Cavalluzzo and Ittner, 2004; Lilian Chan, 2004; Van Dooren, 2005; Gao, 2015; Micheli and Manzoni, 2010; De Waal and Counet, 2009; Greiling, 2006). The literature consequently indicates that our knowledge is still limited when it comes to understanding the use of measures in the public sector as well as the factors that affect the utilization of measures (Carlucci, Schiuma and Sole, 2015; Julnes and Holzer, 2001; Degroff et al., 2010; Ohemeng, 2009; Taylor, 2011a). Additionally, researches on whether the utilization of PM would mean better performance and outcomes, particularly in developing countries, remain thin with evidence of both success and failure to integrate measurement into public management (Elg and Kollberg, 2009; Goh, 2012; Gao, 2009).

This study provides a structured discussion about the issues encountered when adopting and using PMs in public organizations. The study builds upon the findings of a systematic review of the literature and an empirical research to draw some conclusions about the main issues relevant to the development and use of measures in the public sector particularly in developing countries. The study concludes by presenting a new theoretical model that should be helpful to disseminate and understand how measurement policies can be adopted and implemented. Policy makers and scholars interested in models of public administration and policy transfer should be encouraged by findings of this research to continue their efforts to develop a suitable theoretical model for utilization of PMS outside the private sector.

Organization of the paper

This thesis presents the undertaken research throughout the DBA programme which is structured in seven chapters that cover the following:

Chapter 1 begins with a general overview of the research process. It states the research rationale and motivation, and general background or context to the study.

Chapter 2 introduces the Scoping Study which is an early general overview of the literature about different challenges and problems associated with the implementation of performance measurement systems in public organizations. The study highlights possible effects on the overall performance in order to identify suitable questions for an extensive systematic review of the literature.

Chapter 3 provides a general overview about the overall philosophy and methodology that define this research. The chapter explains how the research unfolded throughout time and provides detailed information about the rationale and logic for each step of the study.

Chapter 4 presents project 1 of the DBA, the Systematic Review, which summarizes the current area of inquiry in a way that inform future researches. The Systematic Review is a thorough form of review that adopt a replicable, scientific and transparent process to assess the existing intellectual territory and to specify a suitable research question(s) in order to further develop the existing body of knowledge (Tranfield, Denyer and Smart, 2003).

Chapter 5 discusses the empirical project of the research which seeks to establish practices of Performance Measurement adoption and use within public organizations in the Republic of Yemen. The study reflects the significant importance of the context and subsequent processes and arrangements that promote or impede the development and use of PM as a reform tool in public organizations.

Chapter 6 builds on the findings of past projects to present a new conceptual model for the utilization of PMS in the public sector. The chapter provides an extensive description of the main building blocks of the conceptual model as well as general overview about the rationale and background of the new model.

Chapter 7 is the final chapter that reviews the overall research findings and conclusions. It begins by providing a summary of the main findings from the Systematic Review and the empirical project. This is followed by a brief presentation of the main theoretical and practical contribution of this research. The chapter ends by a brief discussion about the main limitations of this study and possible research areas in the future.

Overview of the Research Process

Introduction

When I first joined the DBA programme of Cranfield, I had little knowledge on the field of performance measurement and management but I knew that there are many problems related to the performance of public organizations in my country. My initial aim was to dedicate my doctorate to make a comparison between public organizations in Yemen and others in developed countries in order to understand the main differences between the two in terms of the problems that inhabit an effective administration of public programs in Yemen. Given my experience in the public sector and what I learnt from my previous studies I realized that one key element that distinguishes between the two was related to the way that public organizations monitor and determine their performance. I realized that the need for effectively administrating public organizations in the republic of Yemen would largely depend on the way public officials manage and determine their performance. My early readings prior to the program helped me to understand that performance measurement supports decision-making process by providing feedback about targets' realization as well as the accuracy of estimates, and that the potential impact on organizational goals has encouraged many public organizations around the world to adopt and implement these systems. As such, the issue has attracted my attention and I was interested to learn more about how to utilize measurement systems as an effective tool for administrative reform in my organization-the Ministry of Foreign Affairs (MFA).

The following are some of the initial questions that I sought to answer at the time:

- How does the MFA measure performance?
- What is the impact on the organization's ability to achieve objectives?
- What are the main constraints on measuring performance in that environment?

As I joined the DBA program my readings had extended and I started to learn more about: the process by which current performance measures are selected and contribute to specific outcome; the key measures of performance and the engagement of stakeholders in evaluation and decision making; and the factors that limits organizations' ability to benefit from the application of performance measurement and management system.

The Critical Review

The DBA program was structured in ways that proved to be helpful in refining my research by first gaining wider knowledge about the issue of interest and second by narrowing down the point of focus in a way that fit for a doctorate. For example, I used the first assessment of the critical review to look into the field of PMS in general. I started by providing a critical review of 12 scholarly papers in order to understand how performance measures can be used to improve performance. The literature has been chosen from a variety of disciplines to reflect different views and experiences relevant to the application and use of PMS in general whether in public or private organizations. Three of the reviewed papers made their case from a conceptual, theoretical point of view, four employed empirical research to generate/support general frameworks, and the remaining five papers extend on theoretical knowledge by using field studies to relate their findings to literature. The critical review provided some indications about the key determinant of successful implementation of performance measures and possible limitations and difficulties. Examining the theoretical and empirical contributions of reviewed papers was helpful to improve my knowledge and understanding about various aspects of the research interest as well as narrowing down my research question.

The Qualitative Research

I then used the qualitative assessment to further build up my knowledge about the field of PMS outside the private sector. The paper empirically examined some of the issues related to the use and usefulness of performance measures in the public sector including number of aspects related to that context such as: functions of performance measures, characteristics of measures, enabling factors for successful use of measures, and various challenges and problems associated with the use of performance measures. This was based on a small scale research exercise that examines public managers' views about the use and usefulness of performance measures in four different organizations in the Yemeni public sector. The empirical work was helpful in enhancing my knowledge about the topic at the same time opened the possibility for determining gaps in that field. I consequently started to realize that it is possible or rather important to question the applicability of many of the underlying assumptions and findings in the literature in contexts other than that of western countries. For example, the facts that all respondents used the term 'performance evaluation' during interviews to refer to performance measurement process reflected the limited knowledge and experience of managers in

that field. One of the main issues that I noticed as a result is the wide ignorance and non use of measures to determine performance in most organizations.

Additional readings show that there is no straight answer to the way that performance measures should be used in public sector organizations and the literature indicates that the mere existence of performance measures does not necessarily mean that performance will be improved. As such, it suggests that organizations would need to consider many elements prior to and following the implementation of PMS.

The Scoping Study

The scoping study worked to determine relevant questions for a systematic literature review that explore the field of performance measurement within the context of the public sector in order to understand the conditions under which PMS can be successfully deployed and used in public organizations. It provided a broad review of the field of interest and general assessment of the relevance and size of the literature. It also aimed at delimiting and focusing the area of concern into a particular issue or question(s) that qualifies for academic research and investigation. The scoping study, as outlined by Tranfield, Denyer and Smart (2003), provides an early justification of the research area and brief discussion about different relevant perspectives, themes and concepts. The following is a brief review of the main themes that were covered in the scoping study.

Mapping the field

My reading into early papers indicated that the full potential of PMS in the public sector entails radical organizational changes and cannot be successful without effective management of the processes of implementation of the new system. Consequently, the literature has been chosen from a variety of disciplines to reflect different and relevant views about this topic. Three main domains of the literature have been identified for the purpose of this study: change management, public management and PMS. The latter domain is the central area of concern and the public management domain was considered the context that is related to the main field of interest. The intersection between the two focused on generating useful information about the role and the use of PMS in public organizations. This entails the examination of various frameworks and approaches to the design of PMS, the impact from applying these systems on organizational performance and problems and challenges associated with their application. The relevance of change management domain was based on the idea that it can provide some guidance about reasons for failure to change in the public sector and

how change can be implemented successfully. It provided some information about various definitions of change, its drivers and different theories and models of change. However, it was clear upon completion of the scoping study that change management is a huge field and that there are very limited papers that fall on the intersection between the three main domains. Upon discussion with my supervisor, the focus of my research shifted towards the intersection between the two main fields PMS and public management.

PMS in the Public Sector

The scoping study provided a broad overview about the use and usefulness of PM. It revealed that PM can achieve several ends such as: changing behaviors, monitoring goals and policies, holding managers and employees accountable for results, providing basis for better planning and control, and ultimately improve the overall performance and management practices (Jackson, 1988; Gianakis, 2002; Fryer, Antony and Ogden, 2009; Kaplan and Norton, 2001). It was also helpful to better understand the roots of PM in the public sector. It revealed that that the practice of PM was essentially imported from the private sector and driven by management and budgeting initiatives. Management by objectives (MBO), planning programming budgeting systems (PPBS), and productivity and total quality management (TQM) are examples of initiatives that contain elements of performance measurement.

Problems and Difficulties Associated With the Implementation of PMS

However, the scoping study also revealed that some management practices such as measurement are not necessarily suitable or directly transferable from one sector to another. Public and private sectors differs in many ways such as: stakeholders, ownership, meaning of profit, stakeholders' level of commitment, selection of managers and leaders, market exposure, etc. It suggested a need for further research to identify how these features affect the application of PMS in public organizations.

The reviewed papers highlighted some of the main challenges and problems associated with the development of PMS in public organizations. Lack of top management commitment, inadequate training, not having a performance management culture, limited information system, difficulties selecting and interpreting measures, limited decision making authority, internal resistance and people not seeing enough benefit from measurement are some of the factors that were cited in the review and thought to hinder the development of performance measurement in the public sector. However, given that there are relatively very few empirical studies on PMS implementation in the public sector, the scoping study suggested that further research is needed not only to identify problems and challenges but also to identify solutions and ways to overcome difficulties. Further, papers reviewed seem to suggest that the success of PMS implementation is related to the situational variables of each organization. It supported the view that additional research is required to investigate the link between the problems and the type of organizations, to identify generic problems (De Waal and Counet, 2009). Furthermore, the scoping study highlighted the need for understanding the

relationship between factors that affect the development and use of PMS in public organizations.

Perceived Value of measurement

The field of change management was helpful to highlight the significance of the perceived value of change which is part of the theoretical contribution of this study. Bourne et al (2003b) indicates that one of the main compelling factors for progress included the perceived benefits from the PMS. Some papers suggest that conducting a readiness assessment exercise that reviews the perceived benefits and priorities of an intervention is a better approach than the one that is limited to identifying the factors that drive or block implementation regardless of its value (Bourne, 2005; Julnes and Holzer, 2001). This made me realize that managing participants' perception is part of the change management process and that the more change is viewed as desirable and inevitable, the more progress can be achieved towards the goal of change (Dibella, 2007). Changing individual employee behavior lies at the heart of organizational change programs, and to be successful in shaping behavior, performance management systems must achieve acceptance by those being "managed" (Cheng, Dainty and Moore, 2007).

The Dynamic of Political support

Another theoretical contribution of this paper that was affected by the scoping study is related to the dynamic of political support. The scoping study reflected that while top management commitment is an important factor for a successful implementation of PMS, it is not constant or absolute. The paper built on the work of Bourne (2005) to suggest that management commitment might be associated with the priority they assign to a certain project and that the priority might change over time. It signaled a need for learning more about the dynamic of management commitment and how it changes over time.

The Systematic Review

My aim with the systematic review was to understand the conditions under which PM could be employed efficiently in the public sector. Identifying all relevant studies that had been conducted on the adoption and implementation of PMS in the public sector was important to form a clear picture about the topic and to learn more about areas that need further researches.

My primary systematic review question was:

What are the key determinants of successful utilization of PMS in public organizations?

Additional sub-questions included:

- What are the factors that affect the adoption of PMS?
- What are the key characteristics of a successful design?
- What are the factors that affect the use of PMS?

The development of the search string evolved over time based on several search attempts and reflection of results. The following table introduces the main keywords that were used to form the final search string.

Keyword	Rational
Performance	Central topic
And	
Measur * OR manag*	Main area of interest
And	
Public OR Government*	The context of the field of study
And	
Implement* OR Adopt* OR Appl* OR Driver*	The inner focus
Not	
Business* OR industr* OR firm* OR compan* OR Corporat*	Excluding papers about the private sector

Table 1 Final Search String

Given the enormous amount of published work around the field of performance measurement and management, the selection of key words was based on a careful attempt to narrow the result to a manageable number of papers suitable for investigation without losing relevant articles. Pre-determined inclusion-exclusion criteria were important for setting boundaries and focusing the search in a way that corresponded to the purpose of the review. The reviewed papers were related to the two main stages of PMS utilization: adoption and implementation. The systematic review provided additional information about the determinants and factors that affect each level of utilization within the public domain.

The Utilization Issue

The first issue that was noted during the systematic review was that the majority of academic papers do not use the term 'utilization' to refer to the development and use of measures in public organizations. Using the term 'utilization' was influenced by the work of Julnes and Holzer (2001), who define utilization as a process of two stages: adoption (development of measures) and implementation (actual use). Authors note that recognizing these two stages is important to isolate the effects of factors affecting each stage. The majority of papers reviewed do not explicitly recognize these two stages or boundaries of utilization. The review supported the findings of Yang and Hsieh (2007) in that most papers are either interested in adoption or implementation and only few were interested in both. Additionally, the meaning and boundaries of 'adoption' and 'implementation' was not the same for all papers. For example, adoption in some papers were viewed in terms of the development of measures (Julnes and Holzer, 2001; Carlucci, Schiuma and Sole, 2015; Lee and Cho, 2011; Brignall and Modell, 2000) while in some others in terms of the driving factors for applying measures (Van Dooren, 2005; Hawke, 2012; Ohemeng, 2011; Siddiquee, 2010; Yang and Hsieh, 2007). Similarly, implementation could refer to the process of measures' development (Berman and Wang, 2000; Cavalluzzo and Ittner, 2004; Sanger, 2008; De Waal, 2003) or to the actual use of measures for various purposes (Van Dooren, 2005; Julnes and Holzer, 2001; Moynihan, 2004; Yang and Hsieh, 2007). The systematic review reflected the complexity of research in this field and showed the difficulty of drawing conclusions based on comparing and analyzing various studies. Given the need to be mindful of the usages of these terms the systematic review paper viewed utilization in terms of two stages: adoption- viewed in terms of: a) the driving factors for applying measurement (the demand side); b) factors that affect the development of measures (the supply side), and c) implementation/use which is related to the actual use of measures for making decisions.

Drivers for Adoption

Papers in the review indicate that adopting PM in public organizations is based on two main driving factors: an internal desire to improve performance, and pressure from external stakeholders. The work of (Cavalluzzo and Ittner, 2004; Lilian Chan, 2004; Bevanda, Sinković and Currie, 2011; Hawke, 2012; Mimba, Helden and Tillema, 2007; Moynihan, 2004; Taylor, 2011a; Verbeeten, 2008) indicated that PMS is more likely to be adopted when the policy of using the system comes from within the organization as an internal requirement. It showed that internal actual desire for improvement is normally driven by top and middle management based on their assessment and valuation of both current practice and the benefits of the new system. On the other hand, pressure from elected officials, responding to citizen demands for greater accountability, and compliance with certain laws, administrative regulations, or reforms represent the external influences that lead to adoption- in the form of demand for performance measurement (Cavalluzzo and Ittner, 2004; Degroff et al., 2010; Folz, Abdelrazek and Chung, 2009; Torres, Pina and Yetano, 2011; Verheijen and Dobrolyubova, 2007; Julnes and Holzer, 2001).

The studies of (Cavalluzzo and Ittner 2004; Moynihan 2004; Folz, Abdelrazek and Chung, 2009; Yang and Hsieh, 2007; Van Dooren, 2005) provided valuable insights about the relevance of the value of measurement initiatives to adoption. It reflected that stakeholders are mainly motivated by the value of measurement. In other words, their demand for performance measurement is dependent on the value they associate with measurement information. When stakeholders believe that measurement information is useful and does not undermine their power or interest they will be more inclined towards the application and use of measurement and vice versa. The study highlighted that identifying the value of measurement is important to understand which measure fits which purpose and to assess organizational readiness to develop and use measures.

Factors Affecting Supply and Consumption of Measures

The systematic review showed that both adoption and use of PM are affected by the same group of factors. That is not to say that factors affecting both levels of utilization have the same influence at each level. The review suggested that the significance of each group of factors are affected by many variables including the maturity of an organization in relation to the use of measures and the level of understanding of the value of measurement. Three main groups were determined: political factors, technical factors and cultural factors.

Political factors are related to the role of internal and external stakeholders who exercise some control and pressure to achieve certain aims. A clear majority of the reviewed papers highlight the significance of internal and external political support for successful

adoption and implementation of PMS. However, while managerial support and leadership is a prerequisite before and after developing measures, external stakeholders' interest in measurement could lead to measurement but not necessarily to the use of measurement information.

Technical factors are related to the available financial and technical resources for the development and use of measures. Empirical and academic literature acknowledges the importance of technical factors in establishing and retaining PMS in public organizations. Most papers in the review were concerned on the following technical issues: the difficulty to measure performance in the public sector; difficulty to allocate the necessary human and financial resources; and the quality of the measurement design.

The systematic review indicates that cultural factors are related to all elements that define and shape culture in the public sector. Several papers refer to cultural influence with no proper definition of which aspects of culture matter to utilization. The review supported the view that there is a need for providing better definition of cultural factors (Lilian Chan, 2004; Taylor, 2011b; Goh, 2012; Hawke, 2012; Sanger, 2008; Inamdar, Kaplan and Reynolds, 2002). Accordingly, reviewed papers were content analyzed in order to identify the key cultural elements that influence adoption and use of measures. Based on the content analysis of the reviewed papers, the bases of cultural influence are: performance-oriented culture and openness to change and learning. Performance-oriented culture refer to organizational culture in which an organization is keen to define its goals and strategies, focus on end results, empower workers, and seek to collect information pertaining to its performance for the purpose of improvement. The systematic review also revealed that different contexts have their own cultural challenges that are not identical for all organizations.

Key Characteristics of a Successful Design

The quality of the measurement design was another concern of the systematic review. Many papers in the review confirmed that findings of Micheli and Kennerly (2005) in that the majority of PMS designs were imported from the private sector into the public sector under the assumption that what works for the former could work for the latter. However, the literature indicated that there are certain features that are unique to the public sector which need to be considered in order to come up with proper design that fits the needs of public organizations. Based on a content analysis, the review suggested that a successful design is characterized by its ability to: translate organizational strategy and objectives into specific measures or indicators, reflect the needs of stakeholders, and incorporate multidimensional perspectives. It should also be valid and fit for purpose, simple and easy to use, encourage stakeholder participation and dynamic that continue to evolve based on the needs or changes in the environment.

Moving to the Empirical Work

Moving to the empirical study was guided by first my main research interest and second the definition of a research gap according to the systematic review. The systematic review revealed that utilization of performance measurement in the public sector continues to be a work in progress. It supported Goh's (2012) view in that the empirical work remains thin and fragmented, focused only on single issues, and lack an integrative perspective. It consequently unveils the following research agenda:

- Past researches were not explicit in terms of defining the meaning or boundaries of the utilization process. They either focus on adoption story or implementation but rarely on both.
- Future research should be mindful of the dynamic of political support which is subject to change over time. Some papers implied a link between political support and the value of measurement.
- The multiplicity of stakeholders in the public sector is a real challenge. Further empirical researches that examine the influence of different groups of stakeholders should be encouraged to understand the balance between different needs and requirements.
- The fact that the review seemed to suggest that external political pressure can lead to adoption and not to implementation need to be challenged and tested empirically.
- Fewer studies have been interested in cultural factors compared to those concerned with political and technical factors, and even fewer seem to focus on cultural influence in non-western context. Further research should be helpful to determine the relevance of cultural factors and to determine whether measurement initiatives are suitable for the developing countries' culture.
- The majority of papers simply listed factors that affect adoption and use of measures with no proper consideration/description of possible interaction that might exist between these factors.

The Empirical Project

The main aim of the empirical study was to offer an explanation for the slow implementation and disappointing effects of PMS in the public sector within the context of developing countries. The systematic review has shown that despite their potentials to improve performance and add value to public services, systems of performance measurement and management is yet to be a prime component of reform particularly in developing countries. Given the limited research on that context, the study adds to the level of understanding about fundamental issues related to the development and use of PMS in public organizations in non-western countries. Reflecting on the experience of several public organizations in the Republic of Yemen, the project focused on enhancing knowledge about the important conditions for effective utilization of measurement in the public sector.

The study aimed to answer the research question:

What are the patterns of PM usage in the Yemeni public sector?

The study was guided by findings from the systematic review and came up with three additional sub-questions that were developed to fill the gap in the current literature on the use of PM in the public sector:

- How do public managers perceive the use and usefulness of measurement?
- What are the types and sources of demand and supply of Performance Information (PI) in public organizations?
- What are the main barriers to the adoption and use of measures?

Findings

Table 2 provides a brief overview of the main findings of the empirical work. These findings are extensively reported in Chapter 5.

Topic	Description
Public Administration	<ul style="list-style-type: none"> • Traditional administration that is not based on management by results and less flexible to change • No clear policy for measuring performance
Cognitive Barrier	Limited knowledge and understanding of the use and usefulness of measurement
Demand and Supply for measurement information	Insufficient demand for performance information and weak accountability culture
External Influence	Disclosure of performance data is driven by external rather than internal pressure
Managerial Control	<p>Limited discretion given to management to control resources</p> <p>Freedom to report performance in their own way</p>
Main Barriers	<p>Limited financial resources</p> <p>Weak qualifications</p> <p>Difficulty to clarify goals and strategies and translate them to measurable tasks</p> <p>High level resistance to change</p> <p>Weak incentive schemes</p>

Table 2 Empirical Findings

**CHAPTER 2: A LITERATURE REVIEW OF PERFORMANCE
MEASUREMENT SYSTEM IMPLEMENTATION AND USE IN PUBLIC
ORGANIZATIONS**

Scoping Study

October 2010

Introduction

The purpose of this scoping study paper is to determine relevant questions for a systematic literature review to explore the significance of the performance measurement function on the performance of public sector organizations. The paper examines the way in which public organizations embrace or change towards the adoption and implementation of performance measurement systems (PMS) and how that might affect performance in these organizations. The paper is guided by an interest to explore the role of PMS in the public sector and within the context of developing countries. The importance of research within such context stems from the fact that there are very few studies that have been published about this topic from that perspective. In the republic of Yemen for instance, the progress of various government's programs has been poor and the application of PMS has been limited to very few public organizations. The ability to determine wither the application of PMS has risen up to expectation and been able to improve performance is yet to be known. The scoping study provides an opportunity for a brief and critical overview of the literature to identify possible gaps for the proposed research area and its subsequent question(s) to explore. The paper provides an explanation for why the research topic is worth investigation for a doctorate, highlighting some of the possible theories that inform the research area and identifying where a contribution to knowledge can be made.

This paper maps and identifies three main domains of literature, as per Figure 1:

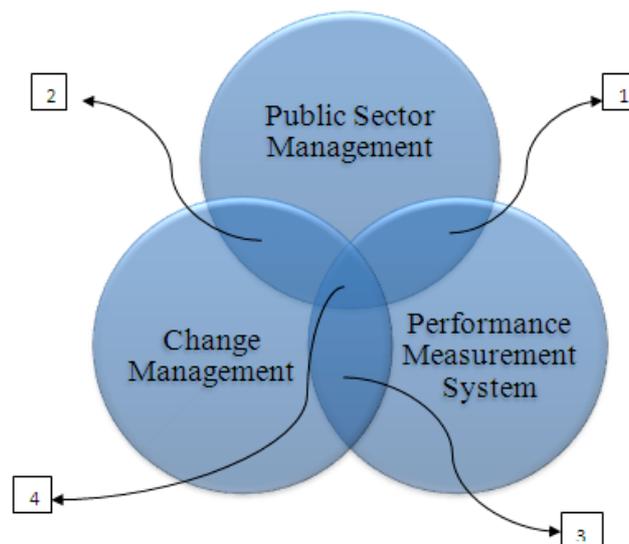


Figure 1 Mapping the Fields

Overlap 1: Measuring performance in public organizations

Overlap 2: Managing change in public organizations

Overlap 3: Implementation of performance measurement system

Overlap 4: Systematic Review Question(s).

General Overview

Measurement of performance in public organizations has gained increasing attention and has been one of the main premises of the New Public Management (NPM) (Bevan and Hood, 2006). According to Phusavat et al. (2009), the success or failure in both organizational and functional performance can be determined by the application of appropriate performance measures. However, compared to its use in the private sector, performance management has been described to be in its infancy in the public sector (Yetano, 2009). This calls for further investigation about the use of PMS in the public sector and how change towards the adoption of PMS influences performance in this sector. That is to say, since the development of performance measurement systems can be divided into the design, implementation and use of performance measures (Bourne et al, 2000) understanding how the use of performance measurement systems may influence management and organizational performance in the public sector entails the examination of a number of aspects:

- General theories of change and principles for effective management of change in public organizations
- Problems and difficulties associated with the implementation of PMS
- Various frameworks (or designs) of performance measurement system
- The result or the impact of the use of PMS on organizational performance in the public sector.

Change Management

The literature on change management has gained large ground and grown rapidly during recent years. The mounting interest in this field can be explained in terms of the costs of both changing and not changing. While it is a matter of life or death for private companies, managing change is crucial to the success or failure of reform in the public sector.

Yet, different writers discussing change management are not necessarily discussing the same thing. In fact, writers seem to have little in common and many view change from different perspectives. For instance, various definitions have been introduced for change management based on the perspectives that describe change and its role within an

organization. For example, Todnem By (2005) defines change management as a process of continuous transfer in direction, structure, and capabilities to meet changing demands of internal and external customers. Melchor (2008, p 14) define change management as “the ability to influence people’s mindsets, culture, attitude and practice to adapt them to a new environment and arrangements.”

Another issue of debate amongst researchers is the cause of change. To Rieley and Clarkson (2001) change is triggered by the way senior management perceives the effectiveness of their organizations. However, this may not explain why some companies decide to change at their peak. Oakland and Tanner (2007) provide further explanation and distinguish between two main drivers for change, internal and external. According to them, internal drivers for change, such as the need to improve quality or efficiency of company products, are considered to be a reflection of other external drivers such as the reaction to customers’ needs or market competition. Similarly, and within the context of public organizations, Melchor (2008) argues that the perception of a problem or dissatisfaction with a current state is the origin of any change initiative. Further, Stewart and Walsh (1992) argue that change or reform initiatives could be instigated by other factors, such as the response to wider social and economic changes.

As with the definition of change management and its drivers, different theories and models of change have emerged. For example, Burnes (1996a) choose to divide theory of change into three main premises: the individual perspective school, the group dynamics school and the open systems school. Whilst the focus of the first school was on influencing the behavior of individuals, the group dynamic school sees change in terms of the ability to influence group values, roles and norms. The open system school, on the other hand, views change as an issue that is related to the organization as a whole by understanding and influencing its subsystems towards better performance.

Todnem By (2005) chooses to describe different approaches of change based on its characteristics and Senior’s categorization of change. Three categories have been introduced: change characterized by the rate of occurrence, change characterized by scale, and change characterized by how it comes about. Different types of change were considered relevant to the first category including incremental, continuous, discontinuous and bumpy change. Two types of incremental changes were cited: smooth and bumpy. While the first referred to change that occurs slowly at a constant and predictable rate, the second witnesses periods of peacefulness that are disrupted by events from time to time. ‘Incremental change’ was used to refer to operational and on-going changes inside an organization while ‘continuous change’ described the strategies that enable organizations to adapt to demands from internal and external sources. ‘Discontinuous change,’ on the other hand, was described in terms of a one-time event that takes the form of a rapid shift from the current position and is triggered by internal problems or external shocks. Todnem By (2005) presents another type of change that is characterized by scale that could include characteristics such as: incremental adjustment, modular transformation and corporate transformation. Whilst modular transformation can be radical and identified as a change that involves major shift in one

department or division, incremental adjustment can be identified by a slow pace of modifications to strategies and various processes within an organization. Corporate transformation is a change that affects the corporation as whole and involves radical alterations in the business strategy.

Two main approaches to change in terms of how it comes about can be identified: planned and emergent. According to Bamford and Forrester (2003), planned change is based on the work of Kurt Lewin that views change as a series of pre-planned steps that aim to move the organization from one fixed state to another by the use of Lewin's action research model and three-step model (unfreezing – moving - refreezing) (Lewin, 1951). The authors claim that the planned change approach has dominated the framework of change management theory for the past 50 years and that environment uncertainty and the assumption that all stakeholders agree on the direction of change paved the way for a new 'emergent approach' to change. The emergent approach views change as a continuous process that seeks to identify the complexity of problems and possible options to deal with them rather than simple approach with detailed plans or projection of change. An emphasis on "bottom-up" rather than "top-down" means of change implementation were thought to be a significant development that distinguished this approach from the planned approach of change.

However, Burnes (1996b) suggests that neither the planned nor emergent model encompasses the full spectrum of change and both suffer from many limitations which, in his view, explain why the contingency approach to change has gained ground. For instance, Burnes claims that the validity of the emergent approach is dependent on whether or not all organizations work in dynamic conditions or should react to an unpredictable environment, which is not the case all the time. Accordingly, Burnes explains that a contingent approach to change is based on the premise that organizational structure and performance are dependent (contingent) on situational variables (environment) and that organizations need to align their structures with such contingencies. This means a shift from one best way of change for all organizations to one best way for each organization. Yet, illustrating with examples, Burnes argues that an organization does not need to adapt to (or be constrained by) the external environment at all times. He introduces his own view of change by arguing that an approach that recognizes that managers do have a choice to influence situational variables is a better approach to change than those that either force change or assume little influence by the external environment.

Dawson (1994, cited in Bourne et al, 2003b), on the other hand, sees change as a process that is influenced by determinants of organizational change and he introduces three major groupings that, in his view, determine change. These include substance of change, politics of change and context of change. Dawson argues that developing a framework for analyzing change is important to explain the shape of change management. To illustrate, Bourne et al (2003b) explain that "processual approaches to change view change in terms of the time frame and influences on the change process rather than as a prescriptive management process to be followed" (p. 257). Accordingly,

both the choice-management approach to change of Burnes (1996a) and the three-stage model of Lewin (1951) were categorized as a processual approach to change because neither provides a prescriptive approach to change but rather a framework for analyzing change. Lastly, Bourne et al (2003b) provide further explanation about strategies for change that could be divided into hard and soft system models. In their view, the hard system approach was developed from the premise that change can be attained by rational and logical analysis of problems to determine clear objectives and the best way for changing. The soft system approach rejects the premise of the hard system model and claims that change is an ongoing journey and process that involves people's reflection on process and results, time to acquire feedback and use of behavioral and social science methodologies to diagnose and resolve problems. Authors conclude that an effective change could be based on both rational objectives and better consideration of human issues involved in the change process.

Implementation of PMS

The relationship between change management and Performance Measurement Systems (PMS) can be viewed from different angles. On the one hand, PMS has become an essential mechanism for making change happen (Bourne et al., 2003a). "They act as both behavioral change tool and enabler of improved organizational performance through being instrumental in driving change" (Cheng, Dainty and Moore, 2007, p. 62). On the other hand, developing and using new PMS often involves fundamental changes that may be threatening in an organization. Therefore, realizing the full potential of PMS usually entails radical organizational changes and cannot be successful without effective management of the processes of implementation and use of the new system. Further, PMS can play an essential role in assessing performance levels before and after change and provide a useful tool to control change programs (Okland and Tanner, 2007).

Bourne et al (2003a) explain that while the literature of change management provides some guidance about how change should be implemented and why change fails, there are relatively very few studies on PMS implementation. The authors worked to identify major factors that determine the success or failure of PMS implementation following action research intervention in ten manufacturing companies. They suggest that out of the many factors that seem to influence success and failure, only six main factors seem to be more relevant. The two main compelling factors for progress included the perceived benefits from the PMS and senior management commitment to the implementation process. The main barriers included effort required for implementation, ease of data accessibility, the consequences of measurement, and being taken over by a new parent company. The authors suggest that many of the barriers are processual issues that can be resolved by a well-designed PMS. A subsequent paper by Bourne et al (2003b) examined the literature of change management to form a better understanding of why many initiatives to implement PMS fail. Their review suggested that while the quality of the design of PMS is important, it does not guarantee success. They suggest that one explanation could be that processes of design as described in the

literature are only partial processes creating the desire for change and providing the first steps for change, but give little guidance on implementation. In their view, this might explain the lack of success but it does not explain why implementation of PMS works in some organizations and not in others. As a result, the authors suggest that managers should not confine themselves to what is published in the performance measurement literature and that they should draw from change management insights when implementing PMS. However, Bourne (2005) explains that while change management literature assumes that senior management commitment is constant and absolute – that it exists or does not exist – findings suggest that these assumptions are incorrect. He explains that management commitment is more related to the priority they assign to a certain project and that the priority might change over time. He indicates that change management literature does not explicitly address this issue, which might be restricting their methodology. As such, Bourne suggests that conducting a readiness assessment exercise that reviews the perceived benefits and priorities of an intervention is a better approach than the one that is limited to identifying the factors that drive or block implementation regardless of its value.

The work of Dibella (2007) corroborates this idea. Dibella maintains that stakeholders' perception of change can be used as an indicator of the relative success of change implementation. He describes two dimensions of change that are of concern to change participants: appeal and likelihood. According to him, the more change is viewed as desirable and inevitable, the more progress can be achieved towards the change goal or outcome. As such, managers need to understand that managing participants' perceptions is part of the change management process.

In their qualitative study on one of the largest construction businesses in the UK to evaluate the impact of PMS implementation, Cheng, Dainty and Moore (2007) conclude that barriers to implementation centered on a lack of senior management endorsement, employees' overt resistance to change, and inadequate training for the implementation process. The authors reach the same conclusion as Bourne et al (2003b) when concluding that the quality of the system is less significant than the processes involved in change. Employee commitment and belief in the benefits of the new system were necessary in order for the PMS to meet objectives. The authors provide a basis for a new implementation framework that identifies different barriers and ways (facilitators) for better change and implementation of new systems.

A recent study by De Waal and Counet (2009) worked to identify the main problems encountered during the implementation and use of PMS. The research was based on a survey that explored the views of 31 academic and practitioner experts in the field of performance management. Based on the survey results, the authors claim that the failure rate of implementing PMS has decreased in the past decade from 70% to 56%. The authors also suggest that out of the 31 problems encountered during the use and implementation of PMS, five main severe problems could be documented: a lack of top management commitment to the process of change; not having a performance management culture; performance measurement getting a low priority or its use being

abandoned after a change of management; management putting low priority on the implementation; and people not seeing enough benefit from performance management. The paper also presents an interesting analysis about the difference in perception between practitioners and academics by ranking responses according to each group. The authors note that eight of the top ten practitioner problems pertain to behavioral issues and two problems were structural in nature. For the academics, only five out of the top ten problems were behavioral while their first problem was structural, which was not even mentioned in the top ten lists by practitioners. According to the authors, this should serve as a 'wake-up call' for researchers to further examine these problems in order to guide practitioners to deal with them. However, the authors seem to limit their research to identifying problems and do not consider the critical factors for successful implementation. Further, it is difficult to distinguish between problems that are related to the implementation of PMS from those related to the use of PMS, due to the fact that the authors do not distinguish between the terms 'use' and 'implementation.' Additionally, the small sample size could affect the validity of the results, making generalization difficult. Authors concede that the research did not investigate the relationships that might exist between problems, which according to them would have been helpful, especially in cases where one problem could be either eliminated or caused by other problem(s).

A study by Julnes and Holzer (2001) examines the factors that affect the utilization of performance measurement in the public sector. Utilization as identified by authors consists of two stages: policy adoption and program implementation. The study, which was based on the results of a national survey of state and local government officials, suggests that utilization of performance measurement should be viewed as a change process that is influenced by rational and political factors. The authors suggest that whilst policy adoption is driven more heavily by factors from rational and technocratic theory, actual implementation is influenced by factors addressed by political and cultural considerations. Findings suggest that implementation of PMS is influenced by rational factors such as the availability of resources and information and by political and culture factors such as the role of external interest groups and risk-taking attitudes.

Another study by Cavalluzzo and Ittner (2004) that examined factors affecting the development, use and perceived benefits of results-oriented performance measures in some government agencies in the US, found that some organizational and technical factors have a significant influence in the use and implementation of PMS in public organizations. Factors such as inadequate training, limited information system, difficulties selecting and interpreting measures, lack of top management commitment, and limited decision making authority were thought to hinder the development of performance measurement in the US Government. Furthermore, findings from their survey suggest that despite the fact that pilot sites have developed performance measures they do not make greater use of the system information. As such, authors conclude that implementation of externally-mandated systems will be symbolic and have little impact on operations, which is consistent with institutional theories. Authors

also suggested that the research was limited to perceptual measures rather than hard measures such as the actual number and frequency of performance measures or actual outcome. They suggest that future investigation “can make significant contribution by examining the actual outcomes associated with the implementation of results-oriented measurement systems” (Cavalluzzo and Ittner, 2004, p. 265).

Public Management

The efficiency and effectiveness of public sector organizations have been under great scrutiny by many stakeholders including the public and decision-makers. The traditional approach to public management, it was thought, need to be replaced with new techniques that improve the value of public products.

Gruening (2001) explains that, the progressive reforms in the United States during the 1920s have witnessed the creation of public administration science, which was based on various principles such as unity of command, hierarchy, specialization and division of work, and delegation of authority. Public managers were responsible for planning, organizing and directing staff and budgets. Following World War II, the classical public administration principles were under great scrutiny which paved the way for the emergence of new public management. Gruening explains that the new discipline drew on the findings of the neo-classics researchers and witnessed the transplant of different approaches from the private sector. He argues that the transplant can be grouped into two management styles: rational and humanistic. The rational style was marked by a heavy reliance on information-gathering as a means to providing the best solutions to management problems, and an emphasis on the objective measurement of performance. Zero-base budgeting, public sector marketing, management by objectives, strategic management and performance measurement are all examples of the rational thinking that dominated during the 1970s. The 1980s, on the other hand, was marked by a shift towards a more organic and humanistic style of management based on the idea that the most successful companies, at least in the US, did not use a rational management style. Examples of the humanistic style of management include approaches such as organizational development, total quality management and culturally-oriented strategic management. However, the literature indicates that the beginning of New Public Management (NPM) can be traced to the late 1970s and early 1980s when it first emerged under the Conservative government in the UK and in the municipalities in the US where the new shift was marked by a movement in many OECD countries towards a NPM approach (Hood, 1995).

There is no single definition of NPM and the term seems to be ambiguous. For Adcroft and Willies (2005), NPM is a reflection of a number of different trends marked by the importation of practices from the private to the public sector, and from one country to another. Others like Fábíán (2010), define NPM as a new form of management that deals with the modernization of public institutions and as distinct from administration science or scientific trend. According to him, NPM does not necessarily mean the

adoption of a corporate management approach to administration but rather improving performance and economic efficiency. To Hood (1991), this new trend to change public management systems is a consequence of international competitiveness and received ideas about institutional design and managerial best practice. It was presented as a framework that is applicable for all organizations at all times and one that seemed to solve all the problems of governments and bring change towards more efficient and effective management. Hood (2000) identifies it as “an international or even global phenomenon, that represents an attempt to correct the shortcomings of traditional public organization in efficiency and service-delivery to citizens, and that one of its central themes is to stress the importance of public managers’ discretionary space or freedom to manage” (Hood, 2000, p. 1).

The literature identifies many principles associated with NPM such as: emphasis on private sector management style (Hood, 1991, 1995, 2000; Adcroft and Willies, 2005), decreasing public sector employment, quality improvement of public administration, creating conditions similar to market competition (Fábián, 2010), privatization, decentralization, the separation of politics and administration (Fábián, 2010; Gruening, 2001), employment of professional managers, increased accountability in resource use, (Hood, 1991), managerialism, downsizing, incentivisation, de-bureaucratisation, competition, emphasis on outcome and results, cost-cutting, efficiency, flexible management (Siddiquee, 2010), budget cuts, contracting out, competition, freedom to manage, accountability of performance (Gruening, 2001), and explicit standards and measures of performance (Hood, 1991; Gruening, 2001). It is clear that with no consensus on the attributes of NPM, many of these principles are more debatable than others. For example, Gruening (2001) explains that the application of NPM is not equated with legal, budgetary and spending constraints at all times as noted by some researchers. Further, the application of NPM entails a wide organizational change that extends to structure, work processes and management principles.

As with any change initiative, NPM has generated much debate about its impact and ability to reform and improve public organizations. The following are some of the criticisms of NPM as identified by Hood (1991):

- NPM has no substance as it did not bring change and was unable to reform public administration radically
- NPM damaged public service and was unable to lower costs
- NPM is a self-serving movement designed to serve the interests of an elite group of ‘new managerialists’

Additionally, Hood highlights the claim of universality which assumes that NPM is applicable for all organizations at all times. Siddiquee (2010) rolls out this idea by suggesting that the value and benefits of NPM cannot be taken for granted, particularly in developing countries. He maintains that advocates of NPM rarely acknowledge challenges and constraints during implementation and that reforms have been limited in many developing countries. Further, De Araújo and Branco (2009) point out that the

application of NPM in several countries was a limited and superficial phenomenon. They suggest that the influence of NPM varied between countries and that its ability to induce change was limited because many countries adopted some elements of the NPM agenda and ignored others.

Managing Change in Public Organizations

In the public sector, various stakeholders, such as citizens and politicians, place many demands on government agencies to improve the value of their products. Reform initiatives are some of the ways that governments respond to the need for change towards better service and performance. In many cases, the reform programs bring about many changes that include structure, work process, management style and even organizational identity. Thus, managing change is an important aspect of any reform program and determines its success or failure. However, managing change in the public sector entails a careful consideration of several pertinent issues, such as the unique characteristics of public organizations, the impact of change or reform based on the NPM policy, and change principles.

Characteristics of Public Organizations

Ostroff (2006) points out that a high-performing public agency resembles a well-run company, as they both have worthy goals, rational processes, accountability bodies and effective leaders. However, author explains that successful change management needs to address obstacles that arise from the profound differences between the two sectors in issues such as purpose, culture and context. For instance, Ostroff explains that, in many cases, the selection of leaders in public organizations is based on aspects other than merit or commitment to reform, such as command of politics and political connections. The short period of time that a public leader spends in his position makes him/her more focused on quick rather than long-term reforms, which is not necessarily to the benefit of the organization. Further, public managers have less flexibility to manage and are faced with various rules that limit their ability to make decisions which, in many cases, can inhibit the process of change management. Finally, Ostroff points out that the multiplicity of stakeholders in the public domain can be a major constraint, as the operations of a public agency are conducted in a 'fish bowl' where initiatives are bound to be unwelcome and met with dissatisfaction by some stakeholders. Boyne (2002) points out that the main conventional distinction between the two sectors is their ownership which is associated with two contrasts: public agencies are controlled by political forces, not market forces; public organizations are funded by taxation, not by direct fees from customers. According to Boyne, the differences between public and private organizations, which dominated public policy and administration, might indicate that management techniques are not suitable to export from one sector to the other. "These variables represent a set of contingencies that require different approaches to management in public agencies and private firms" (Boyne, 2002, p. 118). However, Boyne explains that there is limited evidence to support this claim. Boyne argues that

evidence from 34 empirical studies suggests that only three main differences can be highlighted: “public organizations are more bureaucratic, and public managers are less materialistic and have weaker organizational commitment than their private sector counterparts” (Boyne, 2002, p. 97). Consequently, Boyne argues, rejecting the application of successful practices in the private sector to public organizations does not have solid empirical ground. Nevertheless, Boyne admits that “there is no established body of knowledge on successful management strategies in the private sector that can be easily drawn upon by public agencies” (Boyne, 2002, p. 118). The relative failure of the application of NPM can be viewed in terms of the differences in function and environment between the public and private sectors. For instance, while public organizations are thought to have less market exposure and wider stakeholder interest, the private sector is thought to work with a clearer conception of its aim and customers (Dopson and Stewart, 1990). Finally, Fábíán (2010) indicates that some of the defining features of the private sector such as profit, the relation of supply and demand, continuous expansion and competition cannot be easily transferred to the public sector because both operate with different circumstances. According to Fábíán, “before we integrate market elements and elements of competition into healthcare, we have to examine whether health can be considered as a product” (Fábíán, 2010, p. 52).

Impact of Change

Change is not cheap and does not come without a price. According to Stewart and Walsh (1992), whilst the use of private sector management practices was originally identified to solve the ills of public sector, it generated new challenges and problems. In their view, retaining legitimacy and the difficulty to objectively measure performance influence the decision to import and apply these practices. In effect, the authors indicate that change was introduced to express ideological commitment rather than to solve problems. The language of consumerism, the development of government by contracts, the form of performance management and the use of quasi-markets are some issues drawn from the attempt to apply methods from the private sector to the public sector. Nevertheless, the authors do recognize the strength of changes in the public domain whilst warning against the dangers of applying private practices without question. Consequently, the authors conclude that the management of public services has to be based on an understanding of the unique purposes, conditions and tasks of public organizations. Their paper argues that adopted approaches from the private sector need to be balanced by approaches that recognize the values of the public sector. Yet, they did not provide a clear explanation of the way in which such a balance could be achieved or who should be responsible for achieving it. Adcroft and Willis (2005) lend their support to some of the aforementioned findings by indicating that the importation of private sector principles will commodify services and deprofessionalise public sector workers. Commodification was identified as the transformation towards more commercial relationships with an emphasis on the economic activity of buying and selling. Consequently, transformations through commodification will inevitably

deprofessionalise workers in the sense that it will change “the basis of decision-making such that values become much less important than the rules, regulations and performance measures of the organization” (Adcroft and Willis, 2005, p. 397). Under such conditions, the authors conclude that “it is highly unlikely that the outcomes of NPM and performance measurement will be significant gains for any of the crucial stakeholders in public service provision” (p 398).

Change Principles

The reviewed papers indicate that an effective management of change in the public sector should take into account the following principles:

1. **Change Plan:** Managing change in public organizations should be based on a clear plan that identifies the targets and process of change in addition to the ways to measure the impact of change. According to Barton Cunningham and Kempling (2009), a plan for change is a strategy that describes a series of action steps to secure essential support for change. Fernandez and Rainey (2006) explain that the change strategy offers a direction on how to achieve desired targets, identify problems or obstacles and ways to overcome and manage difficulties. To them, a good change strategy is one that is based on specific policy goals and rests on causal linkage between change initiatives and desired outcome. The clarity of change goals was thought to provide a standard for accountability and to limit the ability of officials to change the policy objectives – thus goals correspond to action. Additionally, the authors indicate that causal linkage between the process and objectives of change should eliminate inconsistent or conflicting directives that might undermine change initiatives. The clarity of change targets is essential for determining the success of a change initiative (Oakland and Tanner, 2007).
2. **Leadership and Senior Management Commitment:** Many studies highlight the importance of effective leadership along with constant and continuous management support to bring about successful change in the public sector. For instance, Fernandez and Rainey (2006) stress the role of champion(s) to manage change successfully. The authors indicate that while some studies stress the importance of having a single change agent to lead transformation, others maintain that change should be supported by a group of individuals who lend legitimacy and marshal all resources to change efforts. These authors conclude that in both cases high level management support and commitment play an essential role in successful change in the public sector. However, they explain that high level management support for change entails a high level of cooperation between politically appointed executives and top-level civil servants. Furthermore, the authors state that whilst leadership continuity and stability is essential for effective transformation, it is one of the main challenges in public organizations because of the frequent turnover of many government

officials. For this reason, the authors conclude that change in the public sector should be and indeed have been led by career civil servants rather than by politicians. However, management commitment is dynamic and the assumption that senior management commitment and support is absolute - exists or does not exist - is misleading as signified by Bourne (2005). Bourne argues that it is more meaningful to assess management commitment in comparative terms over time or stages of the change process. In his opinion, this can be helpful to ensure stability and continuity of the required support. For Barton Cunningham and Kempling (2009), the first principle of implementing change is the formation of a guiding coalition that is responsible for steering the change. The coalition should represent both the formal and informal organization and facilitate change through various actions such as listening to people, gathering information and making adjustments during change process.

3. **Role of Stakeholders:** Winning over stakeholders is essential for any reform program in the public sector. Ostroff (2006) indicates that the public sector has a large number of stakeholders who fall into two groups - external and internal. According to him, identifying external stakeholders and understanding their concerns is a primary task for managing change. Similarly, change cannot be implemented without the full support and cooperation of internal stakeholders. Ostroff suggests that since workers in the public sector often stay at their organizations longer than their leaders, there is great potential to use their knowledge and experience to create and sustain change. Based on his own experience, Ostroff explains that at any given agency, about a quarter of employees are initially receptive to change, a quarter are resistant, and the remaining half are ambivalent. According to him, it is essential to sustain the support of those who are receptive to change whilst attracting others by understanding their concerns and communicating the goals for change. However, the large number of stakeholders in the public sector adds to the complexity and difficulty of firstly identifying all stakeholders and secondly meeting all their competing and sometimes conflicting requirements. Therefore, in the words of Barton Cunningham and Kempling (2009, p. 330) “doing what is ‘right’ in government is a matter of responding to conflicts and negotiating with various interests much more than it is for a corporate executive trying to implement a strategy”.
4. **The Need for Change:** Barton Cunningham and Kempling (2009) argue that to implement any change initiative it is vital to manage resistance and establish the need for change. Responding to significant resistance begins with identifying the positive and negative forces affecting change and ways to overcome resistance, understanding the degree to which the proposed change aligns with the underlying culture and establishing whether there is a need for a cultural shift. The authors explain that “the change literature highlights the fact that people don’t change unless they are compelled to do so” (Barton Cunningham and

Kempling, 2009, p. 334). For instance, they suggest that some of the main successful change efforts are a result of crisis or unusual events which in many cases resulted in radical rather than incremental changes. Fernandez and Rainey (2006) add to that by suggesting that certain challenges, such as crisis and shocks, can be helpful to convince people to accept change and reduce resistance to it. The authors also warn that the high turnover amongst political appointees can be used by civil servants to resist change initiatives unless management ensures their participation in the stages of change which can reduce this kind of resistance. They state that widespread participation in the change process is also helpful to overcoming resistance.

Finally, drawing on the notion of ‘receptivity,’ as used by Butler (2003) to explain how OECD countries are integrating and dealing with change, Melchor (2008) concludes that change or administrative reform in the public sector should have certain attributes in order to be manageable and reach its goals. According to him, change initiatives should be simple, consistent, based on evidence and realistic. Melchor suggests that governments should not underestimate the importance of change while designing and implementing policy reforms. He maintains that effective management of change in the public sector should focus on the future and consider different ways of action to deal with its side-effects.

Performance Measurement

In recent times, an expression like ‘you cannot manage what you cannot measure’ has dominated management literature, particularly pertaining to the private sector. The process of measuring performance was considered a useful tool that could achieve several ends. For instance, performance measurement can be used to: change behaviors (Fryer, Antony and Ogden, 2009); monitor policy goals and hold managers accountable for outcomes (Gianakis, 2002); and provide a basis for planning and control, and improve performance and management practices (Jackson, 1988). Further, “measurement creates focus for the future...and communicates important messages to all organizational units and employees” (Kaplan and Norton, 2001, p.102).

Moreover, since the success or failure of both organizational and functional performance is determined by the type of measures that are used (Cavalluzzo and Ittner, 2004), several frameworks have been developed for measuring performance many of which were originally developed for the private sector. Nevertheless, traditional performance measurement systems that were based on costing and accounting systems have been subject to many criticisms due to many limitations such as short-termism, lack of strategic focus and a tendency to push out other non-financial measures (Bourne et al, 2000; Neely, 1999; Likierman, 1993).

Other approaches, such as the Balanced Scorecard (BSC) introduced by Kaplan and Norton (1992), aimed at providing top management with a comprehensive and balanced view about business. It aimed at integrating financial measures with some other operational measures on customer satisfaction, internal processes and the organizations’

innovation and improvement activities. It was argued that by looking into the business from these four perspectives, managers would align their activities to the main strategy and would be able to decide what is important and what needs to be done to improve performance. This approach gained exceptional recognition to the degree that many believe that it could be applicable to other sectors as well. Kaplan (2001) himself indicates that while the initial focus and application of the BSC was in the private sector, the opportunity to improve the management of non-profits should be greater. However, despite some studies that show the positive impact of using the BSC, including for non-profit sectors (Greatbanks and Tapp, 2007), various complications have arisen regarding its application particularly in the public sector. Kaplan and Norton (2001) recognize that applying their framework in the non-financial sector has some difficulties; On top of which are issues related to the definition of both strategy and customers. The issue becomes more complicated when considering the multiplicity of stakeholders and the share of production or decisions amongst public agencies and actors (De Bruijn, 2007), which is different from that in private companies. Moreover, the framework did not explain how to draw a balance between various and even conflicting needs of different stakeholders. As such, modifying the architecture of the BSC was considered essential for its use in the public sector (Kaplan and Norton, 2001). Some of the proposed modifications however go with other frameworks for measuring and managing performance.

For instance, the Performance Prism approach developed by Neely, Adams and Crowe (2001) assumes that the starting point is the identification of both the organization's stakeholders and their needs. According to authors, the only reason for an organization to have a strategy is to deliver value to some set of stakeholders. They claim that their framework might be a second generation framework that addresses the shortcomings of many traditional measurement frameworks, including the balance scorecard. The framework stresses the need for developing scorecards that are appropriate to the business and consist of five interrelated facets: stakeholders, strategies, processes, capabilities and stakeholders' contribution. Consequently, the authors claimed that a wide variety of organizations, profit and not-for-profit, would be able to relate to and use the framework to improve performance. Nevertheless, the degree of certainty and applicability to a wide range of organizations is limited and yet to be tested further (Micheli and Kennerley, 2005). As with the BSC approach, the term 'stakeholders' is very broad and raises many questions about the balance between competing or even conflicting goals (Propper and Wilson, 2003). There is also the hazard of increasing the number of measures as the number of the organization's stakeholders increases. Thus, measuring the value to or from various stakeholders as well as integrating large number of measures might become problematic particularly in the public sector.

A number of other performance measurement systems were originally developed for the public sector and recognize the differences between public and private organizations. Sotirakou and Zeppou (2003) worked to examine the value of the STAIR model for measuring public sector performance by presenting the preliminary results of a pilot

case study to modernize the Greek civil service. The authors suggest that organizational performance can be enhanced by the steps of the STAIR (strategy, targets, assessment, implementation, results) model, which stress “competence in strategic thinking, strategic acting and strategic measuring, at all levels” (Sotirakou and Zeppou, 2003, p.324). They argue that their model is not limited to measurement but is rather a comprehensive management framework that focuses on organizational success and builds lasting change. However, the model is yet to be tested in other public organizations in different contexts to gauge its applicability and is built on the concept of strategic fit assuming that public organizations have clear, fixed and easily translatable strategies, which might not always be the case.

Others, like Behn (2003), developed a normative argument about performance measures’ selection strategy, based on the identification of the objectives (purposes) of public sector managers. He explains that “it is the nature of the purpose, rather than the nature of the user, that determines which characteristics of those measures will be most helpful” (Behn, 2003, p.593). Behn suggests that a public manager can use performance measures to evaluate, control, budget, motivate, celebrate, promote, learn and improve. As such, a public manager will be unable to decide what should be measured without determining the purpose. He maintains that there is no single performance measure that is right for all eight purposes and introduces selection criteria for each purpose. He then briefly introduces some standards for which performance measures would have a meaning and achieve their purposes. However, as with many other approaches to the selection of PMS, his strategy was not free from limitations. For instance, Behn seem to assume that it is always possible to measure outcome in public organizations and this is not necessarily the case (Cavalluzzo and Ittner, 2004; Greiling, 2006; Kaplan, 2001). And the author does not consider accountability as one of the purposes, regardless of its importance and significance to the public sector (Jackson, 1988; Heinrich, 2002; Likierman, 1993).

Measuring Performance in the Public Sector

The practice of measuring performance in the public sector was initially driven by management and budgeting initiatives that had elements of performance measurement such as: management by objectives (MBO), planning programming budgeting systems (PPBS), productivity and total quality management (TQM) (Julnes, Holzer and Shick, 2009). According to De Bruijn (2002), the idea centered on the formulation of an envisaged performance and the way to measure it.

Yet, many researchers choose to challenge the assumption that performance measurement systems inevitably work for public sector. Indeed, some like Radnor and McGuire (2004) suggest that performance management in the public sector is closer to fiction than to fact. They suggest that performance is about measurement and evaluation not management, systems are diagnostic not interactive, targets baseline is unclear and

neglected, and there is an overall lack of ownership. Accordingly, and building on their previous model, the authors argue that in order for performance management to work in the public sector, understanding of the organizational elements or facets, and the relationship between them, needs to be established. Their paper suggests that these elements are strategy, process, people and system.

Rantanen et al. (2007) worked to identify problems associated with the design and implementation of performance measurement systems in Finnish public sector organizations. A multiplicity of stakeholders and undefined goals were thought to be factors affecting the design process, whilst a lack of ownership of property and poor management skills were considered factors affecting the implementation process. The authors suggest that factors affecting the design phase are related to the nature of public sector organizations, hence are not easy to overcome. They suggest that such factors may be influenced by increasing the use of outsourcing, public-private partnerships and privatization. Factors that influence the implementation process were considered easier to tackle by developing the management practices in these organizations. For example, they may be influenced by increasing personnel education and recruitment from the private sector, as well as by manager selection criteria.

Adcroft and Willis (2005) introduce six main problems associated with the use of performance measures, which they suggest are common amongst public sector organizations. Accordingly, they argue that the increased use of performance measurement and the importation of private sector management principles and practices will have the dual effect of commodifying services and deprofessionalising public sector workers. Commodification was defined in terms of an emphasis on the economic activity of buying and selling and the management activity of performance measurement. However, the authors did not support their claims with empirical evidence nor link them to specific theory or research. They assume that their claims apply to all public organizations within national and international contexts but with no clear basis or evidence to support their level of certainty.

On the other hand, De Bruijn (2002) suggests that performance measurement can fulfill a number of positive functions but may have negative effects that undermine overall performance in public sector organizations. The author explains that professionals would either conform or not conform to the measurement system. In the first case, professionals would be less able to acknowledge the fact that their performance is a multi-value one to be achieved in co-production with others, and perverse effects will occur as the number of functions that performance measurement has to fill increases. In the second case, perverse effects will be produced as professionals manage to shield themselves from the system (game playing). The author then suggests six main strategies to address the perverse effects of performance measurement. However, the introduced strategies remain theoretical and untested in practice. Verbeeten (2008) from his part makes a clear distinction between two strands of performance: quantitative and qualitative. While the former relates to short-term efficiency and quantitative aspects of performance (i.e. use of resources, number of outputs and efficiency), the latter relates

to long-term effectiveness and to both strategic capacity (i.e. long-term effectiveness and innovation) and operational quality (accuracy). The author indicates that the use of performance measures may provide opportunity to increase quantity performance (i.e. use of resources, number of outputs and efficiency) but might not have an impact on quality performance (i.e. long-term effectiveness and innovation). The implication is clear: “Does the easy to measure drive out the more difficult?” (Gray and Jenkins, 1995, p.89). The authors conclude that it appears that the behavioral consequences of performance management practices on public sector managers are as important as the economic consequences, indicating that public sector organizations face a trade-off between achieving quantitative goals and quality goals.

Discussion

Despite their potential to improve organizational performance, many governments have not developed performance measurement systems and fewer use these systems to improve decision making (Julnez and Holzer, 2001). Reasons for this are yet to be understood and literature indicates that there is a deficiency in our knowledge about the success and failure of the implementation of performance measurement systems in the public sector. Yet, papers reviewed suggest that there are a number of different aspects that need to be considered when implementing and using PMS in public organizations.

To start with, few attempts have been made to provide public organizations with PMS tailored explicitly to their needs (Micheli and Kennerley, 2005). Traditionally however, various frameworks and practices have been imported to the public sector from the private sector under the assumption that what works for the latter should inevitably work for the former. However, this is not always the case. Papers reviewed clearly indicate that public sector organizations have distinct features that are different from private companies. This indicates that there is a need for further research to identify proper PMS design, one that recognizes the unique characteristics of the public domain. Additionally, with many papers identifying the differences between public and private organizations, relatively fewer studies have been interested in identifying how that is going to affect the application of private-led practices into public domain.

Additionally, the review indicates that the assumption that PMS will inevitably work for public organizations is flawed. The fact that there are many perverse effects associated with the application of PMS in public organizations raises many questions about their applicability in the public domain. In fact, our knowledge is limited when it comes to investigating the impact of PMS on the performance of public organizations due to very limited number of empirical studies in this field. Thus, further research is needed not only to identify problems and challenges but also to identify solutions and ways to overcome difficulties. Literature also indicates that our knowledge can improve significantly by studying the relationships between different factors in order to gain better information on why and how certain problems happen. Similarly, studying the relationship between problems can be helpful to determine the significance of each as it is possible that there are some adverse effects that can be solved by dealing with the main problems that cause them in the first place. The literature also indicates that while the quality of the application of PMS is essential, it is not sufficient for success. Rather, they represent a partial or first step process for a comprehensive and complete transformation. It is to that end that a research into successful implementation of PMS should draw lessons and insight from change management literature.

Additionally, it is also important to note that our knowledge about factors affecting the implementation of PMS in both public and private sectors is based on the reflection of

practitioners (perception) rather than hard measures which calls for further empirical and comparative studies in this field. Further, challenges and problems associated with the implementation and use of PMS in public organizations did not seem to consider the context on which they occur which could be unique to their respective organizations. Therefore, as De Waal and Counet (2009) argue, additional research is required to investigate the link between the problems and the type of organizations, to identify generic problems. To that end, I also believe it is important to question the applicability of many of the underlying assumptions and findings in contexts other than Western countries. For instance, in an environment that might be characterized by low institutional capacity, limited involvement of stakeholders and high level of informality, such that exist in developing countries (Mimba, Helden and Tillema, 2007), it would be interesting to look into how performance measures may be used effectively in such context and wither the same set of complications that affect the development of PMS, and have been identified by literature, is the same for all countries.

Finally, my review of the papers drives me to think that my field of interest seems to be in line with the contingency approach of change theory. That is, the development of new PMS often involves fundamental changes and the variation in the experience of different organizations indicates that the success of PMS application could be dependent (contingent) on the situational variables facing these organization. However, I understand that this is not conclusive at this stage and that the systematic review should help me to gain better insights about my field of interest and better position my research within the current body of knowledge.

Systematic Review Questions

The literature suggests that there are many challenges and problems associated with the implementation and use of PMS in public organizations. My field of interest is centered on identifying best approaches to implement and use PMS in public organizations, more particularly in the Yemeni government which falls under the classification of developing countries. However, the fact that there are few researches about the success and failure of PMS implementation in both public and private sectors in general, the chance to find enough literature within the context of developing countries seems to be very slim. Alternatively, investigating the gap in determining how public organization can implement and use PMS efficiently, regardless of the context, should provide useful information for future research. The systematic review will be helpful in identifying the critical factors for success and failure of different stages of the development of PMS namely: the designing, implementation and use of these systems. Accordingly, knowledge gained from the systematic review should then be put into test by further empirical studies to determine its validity. Consequently, my systematic review questions are:

- What are the key determinants of successful design and implementation of performance measurement systems in public organizations?
- How should PMS be used to enhance performance in government agencies?

Scoping Study Summary Paper

Despite their potential to improve organizational performance, many governments have not developed performance measurement systems (PMSs) and fewer use these systems to improve decision making (Julnez and Holzer, 2001). Reasons for this are yet to be understood and literature indicates that there is a deficiency in our knowledge about the success and failure of the implementation of performance measurement systems in the public sector. The purpose of the scoping study paper is to determine relevant questions for a systematic literature review that explore the field of performance measurement within the context of the public sector. That is to say, examining the literature in order to understand the conditions under which PMS can be successfully used in public organizations. Consequently, the literature has been chosen from a variety of disciplines to reflect different and relevant views about this topic. Three main domains of the literature have been identified for the purpose of this study. While the first domain relates to the central issue of PMS the public management domain was considered the context that is related to the main field of interest. The intersection between the two domains focused on generating useful information about the role and the use of PMS in public organizations. This entails the examination of various frameworks and approaches to the design of PMS, the impact from applying these systems on organizational performance and problems and challenges associated with their application.

Yet, realizing the full potential of PMS usually entails radical organizational changes and cannot be successful without effective management of the processes of implementation and use of the new system. The relevance of the third domain of change management can be better explained by the work of Bourne et al. (2003b) who argues that both the quality of PMS design and application are “all partial processes in that they create the desire for change and provide the first steps for change, but give little guidance on implementation”. In other words, since there are relatively very few studies on PMS implementation the literature of change management thought to provide some guidance about how change should be implemented and why change fails. Consequently, while recognizing that the field of change management is a huge domain the relationship between this domain and the two other domains can be viewed in different ways. For example, while the intersection between change management and PMS focused on identifying the contingent factors that inform both the design and the implementation of these systems, the intersection between change management domain and public management focused on identifying the principles of change that is more applicable to public agencies and recognize their unique characteristics. The review of the intersection between the three main domains drives me to think that my field of interest seems to be in line with the contingency approach of change theory which according to Burnes (1996b) is a move from the one best way for all to the one best way for each. That is to say that the development of new PMS often involves fundamental changes and the variation in the experience of different organizations indicates that the

success of PMS application could be dependent (contingent) on the situational variables facing these organization. However, it should be noted that this is not conclusive at this stage and that the systematic review should be helpful to gain better insights about my field of interest and better position my research within the current body of knowledge.

The following section provides further account and general overview about some of the issues that were discussed in the scoping study paper and relates to the intersection between the three main domains.

The Design and Implementation of PMS

Papers reviewed suggest that there are a number of different aspects that need to be considered when designing and implementing PMS in public organizations. For instance, the literature indicates that few attempts have been made to provide public organizations with PMS tailored explicitly to their needs. Instead, various frameworks and practices have been traditionally imported to the public sector from the private one under the assumption that what works for the latter should inevitably work for the former. Papers reviewed refute this assumption and clearly indicate that public sector organizations have distinct features that are different from private companies which inevitably affect the design of PMS. Additionally, none of the PMS frameworks introduced in the review managed to deal with the difficulty to identify all stakeholders that a public organization has. According to Micheli and Kennerley (2005), ‘the complexity this issue raises will probably be one of the hardest challenges to face’. Further, the issue becomes more complicated when considering the multiplicity of stakeholders and the share of both production and decisions amongst public agencies and actors. Consequently, the likely trade-off between competing and possibly conflicting goals of various stakeholders in a public organization did not seem to be recognized or clearly explained by these frameworks. Further, many of the well known frameworks to date suffer from an unmistakable weakness in terms of their inability to identify cause-and-effect relationships between elements in these models. Further, while most of these frameworks provide a general overview about measuring performance focusing on what to measure they provide little guidance on how to actually measure or implement PMS in public organizations. Our knowledge is still limited in terms of deciding about the best approach to adopt in public organization. This indicates that there is a need for further research to identify proper PMS design that recognizes the unique characteristics of the public domain, identify the relations between different factors and balance between them, provide clear guidance into implementation process, and take into consideration many of the limitations of previous models.

The review indicates that the assumption that PMS will inevitably work for public organizations is flawed. The fact that there are many perverse effects associated with the application of PMS in public organizations raises many questions about both the design and the quality of implementation as well as factors affecting the two. However, our knowledge is limited when it comes to investigating the success and failure of the

application of these systems in public organizations due to very limited number of empirical studies in this field.

Papers reviewed also indicate that top management commitment is an important factor for a successful implementation of PMS. This said to be true in private as well as public organizations. Yet, it is worth to note that most of the papers reviewed viewed management commitment as a constant and absolute factor - that it exist or does not exist- which according to Bourne (2005) is misleading. For instance, Bourne explains that management commitment is more related to the priority they assign to a certain project and that the priority might change over time. That's why our knowledge should significantly increase by learning more about the dynamic of management commitment and how it changes over time. Additionally, except for this factor, papers reviewed seem to suggest that there is no consensus about factors affecting the implementation of PMS whither in public or private sectors. The variation in the experience of different organizations indicates that the success of PMS implementation could be dependent (contingent) on the situational variables facing these organizations which go in line with the contingency theory of change. As such, future research should not only work to identify factors affecting the implementation process but also identifying generic factors from those unique to their respective organizations. Further, additional researches are required to investigate the link between different factors so to identify those can be easily eliminated by addressing the more important ones. Moreover, it is also important to note that our knowledge about factors affecting the implementation of PMS in both public and private sectors is based on the reflection of practitioners (perception) rather than hard measures which calls for further empirical and comparative studies in this field. Finally, investigating the gap in determining how public organization can implement and use PMS efficiently should be helpful in identifying the critical factors for success and failure of different stages of the development of PMS namely designing and implementation. Accordingly, knowledge gained from the systematic review should then be put into test by further empirical studies to determine its validity. Consequently, an appropriate systematic review questions could be:

What are the key determinants of successful use of PMS in public organizations?

Few sub-questions emerge from that as follows:

- *What are the factors that affect the adoption of PMS in public organization?*
- *What are the key characteristics of a successful design?*
- *What are the enabling factors for successful implementation?*
- *What are the blocking factors?*

CHAPTER 3: RESEARCH METHODOLOGY

Research Philosophy

In social science, researchers need to think extensively on what constitutes a valid research and how to acquire knowledge about the research area. The philosophical and theoretical stance of the researcher is important as it interrelates with the research methodology and used methods for data collection. As outlined by Gray (2004), while the choice of methods is influenced by the research methodology, this methodology, in turn, is dependent on the ontological and epistemological perspective of the researcher.

“Ontology is the study of being, that is, the nature of existence. While ontology embodies understanding what is, epistemology tries to understand what it means to know. Epistemology provides a philosophical background for deciding what kinds of knowledge are legitimate and adequate”

(Gray, 2004, p.16).

Different philosophical perspectives are based on basic assumptions about the world, how it is perceived and how to be able to understand it. O’leary (2004) provides a useful way to what she calls ‘wading’ through various methods. According to her, having some understanding of the discipline can be helpful to ground the research approach and the first question that should be asked relates to how one understands the world. Such question is related to the set of assumptions that define an intellectual understanding of how the world operates and the way knowledge is produced. The next question is about learning what methodological approach will best answer the research question. The answer to that question involves the set of strategies and procedures that would be used to gather and analyze data. The methodology is often driven by paradigm and the discipline, and covers range of different approaches. The third and fourth questions are related to the techniques that could be used to collect and identify different types of data. O’leary (2004) also explains that given the plethora of literature on quantitative and qualitative methods, both strategies refer to the type of data and not to the type of method. The last question will be, of course, about the way researchers analyze or make sense of their data. While quantitative data might be analyzed statistically, qualitative data, according to her, is better to be analyzed thematically.

The most influential philosophical perspectives are positivism and interpretivism. Positivism claims that social world exists externally to the researcher - there is an objective reality- so the research would be primarily based on discovering the objective truth. The main characteristics of positivism are that researchers must be remote and independent from participants or the subject area, linear relationship between causes and effects (no causes without effects and vice versa), concepts can be operationalized meaningfully making it possible to measure facts quantitatively, and possibility to

generalize observations free from situational and contextual constraints (Bryman and Bell, 2007; Gray, 2004).

Many of the outlined characteristics of the positivist approach did not seem to be applicable to this study. For example, coming from a practice background and being a civil servant who would deal with other civil servants in relation to the current investigation, it is hard to expect that the researcher can easily be isolated from the research area. Moreover, reflecting on my early research (qualitative assignment) and the systematic review, it seems that there is no objective reality to be discovered in the true sense of the meaning. As indicated in previous sections, not only the literature indicates that there is still limited work on the application of PMS in the public sector, it also points that researchers in this field are still confused or not able to agree on the same meaning of different terms such as performance measurement versus performance management, adoption versus implementation, implementation versus use of measures, and successful utilization. In view of all that, it appears to me that concepts are not necessarily operational- carrying the same meaning- which makes it difficult to use quantitative measures or to generalize findings.

Consequently, this research was conducted from an interpretive perspective using a multiple case analysis which is suitable to the interpretive approach (Yin, 2002). The decision to use an interpretive method was pragmatic based on past experience and previous knowledge of the nature of the issue being investigated. It was also based on my understanding about the significance of the relevant context - public organizations in developing countries in this case. The systematic review helped me realize that there seems to be a dominance of positivism as the underlying approach in relation to studying the field of PMS in the public sector. Most noteworthy and core papers were quantitative in nature based on small or large scale surveys for data collection. The dominance of New Public Management doctrines, including that of performance measurement, in western countries in particular is one possible explanation for choosing that approach. However, in unfamiliar conditions such as that in developing countries where there is little knowledge to determine specific construct, variables and relationships to be tested descriptive and inductive qualitative research strategy might be more helpful (Partington, 2002). The conclusion I reached was that if early results from the systematic review and my small qualitative assignment research suggest that there is limited knowledge about PM and its use in the public sector within the context of the study it might not be sensible to use quantitative methods in my empirical projects. This meant that in order to investigate and explore the potentials or challenges for successful utilization of measurement in public organizations it was important to first develop an understanding of the views of key stakeholders about this matter which is best obtained through interpretive methods. As outlined by Gerring (2007), an interpretive quest of the social science recognizes that actors do influence and get influenced by their social reality which is why it is important to understand behavior from the actors' point of view. Reality is reflected by interpreting what key actors are saying about their experience and how they make sense of their world.

Having said that, I must probably also mention that I was more open for the use of multiple methods in my research. I'm with the view that the dominance of a single set of philosophical approaches might be restrictive and there is much to be gained from greater plurality of methodologies. In fact, one of my early intentions when I started the Systematic Review project was to build on the first empirical project to investigate the issue further possibly by increasing the sample size using quantitative research strategy. The point that was made at the time is that quantitative data may confirm or refute earlier findings and could be helpful for generalization purposes. However, civil conflicts erupted in 2011 and subsequent civil war in the country paralyzed the public sector and consequently made it difficult to pursue that option.

Research Strategy and Design

The research strategy and design were adopted from a pragmatic point of view to accommodate with the research questions, research philosophy and changes in the context of this study. Recalling that the empirical project was exploratory in nature with the purpose of determining the patterns of performance measurement application in the Yemeni public sector, my initial aim was to focus on as few organizations as possible. This is based on the notion that the fewer cases there is the more intensively they are studied (Gerring, 2007). However, as I started my early arrangements for the research I became more aware about the impact of recent political conditions following the Arab Spring on public organizations in Yemen and by effect on my research. For example, political tension was high to the point that some public officials who have already agreed on participating in this study unexpectedly changed their minds. Many others who were still willing to participate did not seem to be happy to share performance data or invite me to attend performance review meetings. Further, many public facilities needed to close or relocate due to arm conflicts, mass demonstrations or organized sit-ins in major cities in the country. Evidently at that point, it was no longer possible to investigate cases intensively. However, I was still certain that I would be able to gain access to individual officials and senior managers in other public organizations. Since my aim was to obtain and examines the views of several public officials on the actual and potential role of PM to support reform and improve performance in the public domain, the advent of the NDC was seen as hard to miss opportunity to meet with as many decision makers in the government as possible.

Taking that into consideration and noting the abovementioned points about the research philosophy in addition to the difficulty to monitor or rather obtain reliable and quantifiable performance data in the Yemeni context, the empirical project was carried out using qualitative approach based on a cross-case comparison of several public organizations in the republic of Yemen.

Yin provides the following definition of a case study:

“A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident”

(Yin, 2002, p.17)

According to Gray (2004), case studies are most useful when it comes to exploring and understanding many themes and subjects from a much more focused range of people, organizations or contexts. Such design explores subjects and issues where relationships may be ambiguous or uncertain and can be used for:

“a wide variety of issues, including the evaluation of training programs (a common subject), organizational performance, project design and implementation, policy analysis and relationships between different sectors of an organization or between organizations”

(Gray, 2004, p.123).

Approaching this study qualitatively was also influenced by several papers in the field of PMS. For example, whereas Van Dooren (2005) suggests that qualitative approach may be more helpful to identify enabling and impeding factors for measurement in public organizations, Rantanen et al. (2007) indicate that it is known to have greater potential to enhance knowledge about performance improvement. Moynihan (2004) points out that the field of performance management research is guilty of failing to learn how key actors understand their context in which they make decisions. He explains that taking into consideration the perspective of main actors provides a richer picture about policy implementation and why some reform outcomes occur while some others don't. On the same principle, Torres, Pina and Yetano (2011) define the case study as a qualitative research approach in which the collection and analysis of data is the primary instrument of research. Authors indicate that the process is exploratory and inductive, and is most useful when there is little knowledge about variables and theoretical foundation.

Limitations of Case Study Research

One of the most noted limitations to the case study approach are the limited generalizability of findings. As outlined by Gerring (2007), unit homogeneity across the sample and the population is not assured which is why cases are not perfectly representative of that population. However, Yin (2002) provides a plausible argument by arguing that even scientific experiments cannot lead to generalization if it is solely based on a single test. He indicates that if scientific inquiries need to be replicated

before generalizing findings, the same approach applies to case study's research. While Yin indicates that the analytical benefits from having multiple cases may be substantial, he also maintains that case studies might best be generalized to theoretical propositions instead of certain populations or universes. According to him, case study's inquiry benefits from prior development of theoretical propositions to guide data collection and analysis. The goal will be, consequently, to expand and generalize theories and not to enumerate frequencies. Accordingly, this research used the first project (the Systematic Review) to form preliminary theoretical propositions that was then consolidated and refined by empirical data (for further details please visit Chapter 5).

Another limitation is the noted difficulty to replicate case studies. Replication is pursued in the interest of trying to uphold external validity which is important for generalization purposes. The detailed description of data collection process and the development of an interview schedule with predetermined set of open-ended questions, provided by this study, might be helpful for replication purposes. It can be considered as one replication logic in multiple-case studies (Yin, 2002). However, it must be noted that this study was not seeking a direct replication because for many reasons amongst which is the difficulty to duplicate the exact conditions that govern the research process.

Regardless of these limitations, multiple case study researches have one more advantage in relation to practice in my opinion. This study suggests that results of cross-case comparison would receive more appreciation from practitioners than other type of methodologies. This is because practitioners tend to be either skeptic about or unable to relate to quantitative data. In contrast, managers who might review the outcome of a cross-case research might evaluate the methodology favorably since it is based on the views of other managers in the same field. Accordingly, it is assumed that the probability of implementing the recommendations of this type of research might be higher than other types of researches that employ different approaches.

Data Collection

The conditionality of time and resources is an important factor that could influence the research by focusing on using people own words and interpretation to describe the level of significance of the issues at hand (Mikkelsen, 2005). And, since collecting data through extensive discussions with key stakeholders has better potential to enhance understanding about performance improvement in the public sector (Sotirakou and Zeppou 2006), in-depth interviews were considered the most suitable method for data collection. Bearing in mind that case research can be based on purposive and theoretical sampling as opposed to the principle of randomness in statistical research (Rantanen et al. 2007), the choice of cases was initially purposeful and based on availability. My initial aim was not to be selective in terms of which organization or manager should be

part of the investigation, but random selection was no longer an option for many reasons- some are outlined in the previous section- on top of which the difficult political atmosphere in the country at the time of the study. However, even though the empirical work adopted convenience and snow ball sampling, respondents were selected based on their knowledge and experience in managing their organizations' performance (Chapter 5 provides further accounts for the justification and rational of the used methods).

Accordingly, I developed an interview schedule with a list of open-ended questions to explore, in as much details as possible, the views of several public officials on the actual and potential role of PM to support and improve performance in the public sector. The interview schedule first focused on learning whether public officials are familiar with the concept of measurement and second on eliciting how they manage performance in general. Open-ended questions was helpful not only to determine what types of measures or indicators used to determine performance but equally important to learn more about possible barriers for adoption and implementation of measurement in that context. While the first stage of data collection involved live interviews with 12 senior and middle level managers in five different public organizations, the second stage involved 16 interviews with some public officials who participated in the Yemeni National Dialogue Conference (NDC). What is worth mentioning here is that the course of interviews was not equal all along- might differ from one interview to another. For instance, while the interview protocol included predefined open-ended questions, the specific order and phrasing of these questions were largely dependent on the context of each interview- according to interviewees' interaction and retort with questions. The third and final round of data collection was based on two main focus group meetings. This was conducted following each presentation that I made before around 120 members of the two main committees of the NDC: the 'State Building Committee' and 'Comprehensive Development Committee'. My role during these meetings was a mediator that asks questions and encourages further debate.

Data Analysis

All interviews were conducted in Arabic, the native language of both interviewer and interviewees. Consequently, the analysis was also conducted in the same language in order to maintain the integrity of data and avoid any possible loss of meaning.

The analytical strategy was based on the guidelines provided by Miles and Huberman (1994) which can be summarized into the following:

- Examining transcribed material more than once and taking notes
- importing transcribed material and field notes into NVivo for coding
- creating a summary document form for each interview
- initial coding- the process by which data are broken down into component parts with particular names

- continuous process of review and refinement of textual data to assess and classify important and relevant parts
- generating a tree of all nodes
- reflecting on similarities and discrepancies from the literature
- The arrangement of parents and child nodes into headings for the write up.

The entire process was conducted with the outmost attention to be careful in conveying the meaning and the outcome of the analysis. Further details regarding data collection and the analytical process can be reviewed in the research methods' section of Chapter(5)

Ethical Considerations

The Cranfield School of Management requires researchers to apply for ethical approval prior to the collection of data. The School guideline indicates that there are ethical implications for obtaining data from people or for using data that is not in the public domain. It has an online process whereby researchers and students can apply for ethical approval for each assignment or project before embarking on the research. Consequently, permission was received to carry on with data collection upon the completion and submission of online ethics forms.

With regards to the confidentiality of people involved in this research and their respective organizations, it should be noted that it is strictly maintained and will continue post this research. Consequently, all information about participants' names, positions, or place of work has been altered to uphold that commitment.

Chapter Summary

This chapter provided a general overview about the approach to the research including the overall philosophy and methodology. It detailed the rationale and logic for each step of the study in a way that makes it possible for readers to understand the orientation of this thesis. The chapter shows that this research was conducted from an interpretive perspective using a multiple case analysis. It shows that the decision to use descriptive and inductive qualitative approach was pragmatic based on previous knowledge of the nature of the issue being investigated and on understanding the significance of the relevant context- public organizations in developing countries- which is under-researched area where there is limited knowledge to determine specific constructs, variables and relationships. In this chapter, I reflected that it is my belief that in an unfamiliar ground there is no objective reality- which is why reality is best reflected by interpreting what key actors have to say about their experience and how they make sense of the world around them. The chapter explained the processes for data collection and analysis and how the research unfolded throughout time in order to strengthen the validity of results and ensure that the research is replicable.

The next two chapters will include specific sections to describe the detailed research's choices for a systematic review of the literature and for an empirical piece of work. The methodology section in the Systematic Review Chapter provides a clear description about the processes that were used to select articles relevant to the research area. It contains a description about the definition and purpose of the Systematic Review, relevant questions for the review, and the principles that guide the entire process of the review. The research methods section in Chapter 5 explains the rationale and underlying perspectives that have been used for the empirical project. It provides a detailed description about the nature of the research area and justifications for the use of descriptive and inductive qualitative research strategy. It shows that the limited knowledge about variables and theoretical foundation along with the conditionality of time and resources are some of the main issues that influenced the design of that research.

**CHAPTER 4: PERFORMANCE MEASUREMENT UTILIZATION IN THE
PUBLIC SECTOR: A SYSTEMATIC REVIEW**

1. Overview

Measurement of performance in public organizations has gained increasing attention and has been one of the main premises of the New Public Management (NPM) (Torres, Pina and Yetano, 2011). Many public organizations that seek to improve efficiency, minimize waste and provide value for money services are showing more interest in measuring and managing their performance. The potential impact on organizational goals has encouraged the use of measures and indicators as an effective tool to influence behaviors and improve performance. Today, many reform programs are based on the principle that public organizations are required to measure and report on their own performance to upper levels of authority. Various legislations and administrative mandates in many countries have been developed under the assumption that mandated reporting of performance information can increase public managers' accountability as well as efficiency and effectiveness of the public work (Cavalluzzo and Ittner, 2004). These mandates require and create processes for the development of measures (Sanger, 2008). However, despite their potential to improve organizational performance, many governments have not developed performance measurement systems (PMS) and fewer use these systems to improve decision making (Julnes and Holzer, 2001; Sole, 2009). Reasons for this are yet to be understood and the literature indicates that there is a deficiency in our knowledge about the success and failure of PMS utilization in the public sector (Moynihan, 2004; Sanger, 2008; Mansor et al., 2012). In fact, while the field of performance measurement exhibits a clear growth over the past decades the number of empirical studies is small compared to the overall literature on this topic (Taylor, 2011a). Hence, even for those organizations that have adopted PMS the effectiveness of these systems is still unclear (Mansor et al., 2012). Furthermore, considerable differences between organizations do exist and should be taken into consideration before reaching a conclusion about the organizational and contextual factors that facilitate or impede the utilization of measurement in the public sector (Van Dooren, 2005).

This paper provides an examination of the literature in order to understand how the utilization of performance measures can be productive in public organizations. That is to say, examining the literature in order to understand the conditions under which PMS can be adopted and used successfully in the public sector. It aims to introduce an overview about different challenges and problems associated with the use of performance measurement and main principles for an effective utilization of measures in the public sector.

I have split this paper into two main parts: the first deals with the process of the systematic review (methodology) and the second deals with the findings within each area of the review. In the first part, I began by setting out a brief background about the objectives of the review, down to a detailed description of the systematic review

process. This includes a description about the review panel and protocol, selected databases, search strings, assessment and quality criteria, and the detailed search process. The methodology section ends by providing a descriptive and thematic analysis of the findings as well as relevant limitations.

In the second part of the review, I provide further explanation about each area of the literature by organizing the findings from the systematic review according to the main research question and sub-questions and by providing my own synthesizes. The second section of this part entails a summary of findings and gaps in the field of interest as well as suggestions for further empirical research.

2. Methodology

2.1 Introduction

The main objective of the methodology part is to give readers a clear explanation about the processes that were used to ultimately select relevant articles for the systematic review. The chapter will begin by introducing a small description about the objectives of the Systematic Review (SR) and the relevant review questions that emerge from the previous scoping study. This will be followed by a detailed description about different elements of the systematic review methodology, such as the protocol review and the process of the systematic review. I will discuss and explain the principles that were used to initially identify: relevant key words, relevant databases, inclusion/exclusion criteria, the way and stages in which the search was conducted, and quality assessment of papers.

2.2 Objectives

The objectives of the systematic review were identified upon an early investigation of the literature in the scoping study (a precursor to the systematic review). During that stage, the researcher has identified a gap in the field of Performance Measurement (PM) within the context of public organizations. The purpose, as outlined by the scoping study, was to conduct a thorough and critical review of the vast body of literature in the field of performance measurement, in order to understand the conditions under which a Performance Measurement System (PMS) can be employed successfully in public organizations. The following primary question was developed:

What are the key determinants of successful utilization of PMS in public organizations?

Within the context of the public domain, further sub-questions have emerged from the original one, which provide further explanation about the objectives of this study:

- What are the factors that affect the adoption of PMS?
- What are the key characteristics of a successful design?
- What are the factors that affect the use of PMS?

2.3 Systematic Review Back Ground

The literature review is one of the main parts of any research projects. Tranfield, Denyer and Smart (2003) explain that the literature review provides best evidence for informing policy and practices in any discipline. According to them, the aim of the review is not limited to mapping and evaluating the current intellectual territory, but equally important to develop the existing body of knowledge by identifying suitable questions for further inquiry. However, authors warn of a bias of some researchers and a lack of rigor characterizing some of these reviews. They explain that an approach based on narrative description might have its negative implications on the significance of the research work over the provision of guidance and insights for future policies and interventions. This, according to authors, calls for more scientific and systematic review that should work to provide collective insights through theoretical synthesis into fields and sub-fields.

Systematic review, according to Rousseau, Manning and Denyer (2008), means comprehensive accumulation, transparent analysis, and reflective interpretation of all empirical studies pertinent to a specific question. This means that the systematic research synthesis should be able to evaluate claims while at the same time recognizing boundaries, omissions, and untested assumptions. Put simply in their own words, it separates replicable findings from noise. For the NHS Centre for Reviews and Dissemination (2001), a systematic review is a review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant primary research, and to extract and analyze data from the studies that are included in the review.

In response to criticisms of the traditional approach for reviewing the literature, and in order to provide more rigor and enhance the quality of the research work, the School of Management in Cranfield University has adopted the systematic review process. This research follows the same approach and works to examine and review the literature in a systematic and transparent way as outlined by the systematic review process in the following section.

2.4 Process of the Systematic Review

The systematic review was based on Cranfield systematic review process as stated in the previous section. Figure (1) provides further explanation about the stages of the review and various steps for each stage. The first stage was concerned with planning of the review during which a proposal was prepared followed by a protocol review that identified the aim of the review, the review questions, the search strategy, the search strings, exclusion-inclusion criteria, and sources of information or database that will be used. Consultations and approval of my panel were deemed necessary prior the start up

with stage 2 of the review. The following sections provide further discussion on some of these stages and steps.

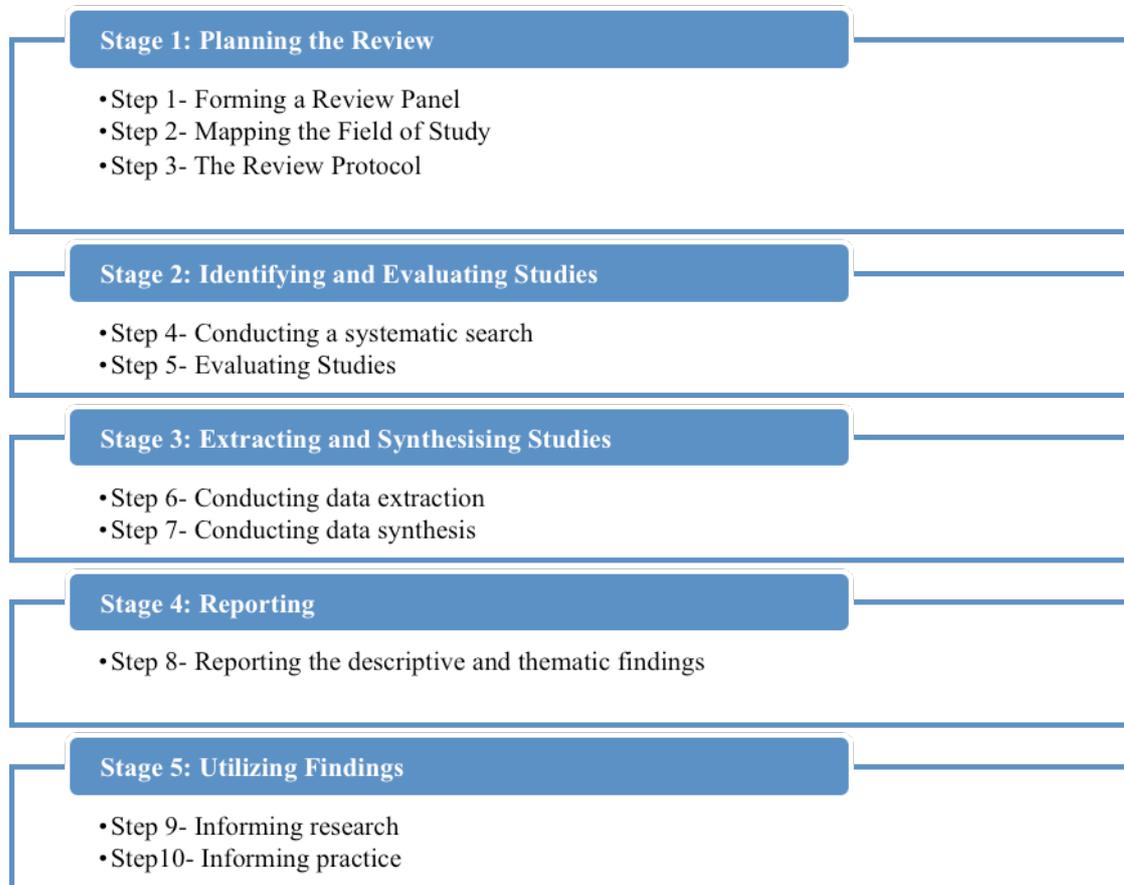


Figure 2 Cranfield Systematic Review Process

2.5 Planning the Review

2.5.1 *The Review Protocol*

According to Tranfield, Denyer and Smart (2003), a review protocol is a plan that provides explicit descriptions of the steps to be taken. It contains essential information that is vital for objectivity such as: review questions, population or focus of the study, list of review panel, search strategy, and criteria for including relevant articles and excluding irrelevant ones. These are discussed in more detailed in following sections.

2.5.2 Forming a Review Panel

Table (3) provides a list of the names and occupation of the SR panel members. Panel members were selected based on their knowledge and experience pertaining to the field of study and the process of conducting the systematic review. The selection was mainly based on consultations with my supervisor and with some help from my cohort leader and friends. The aim was to have a diverse range of expertise in different fields in order to enrich the discussion and add to the value of the SR quality. As can be seen from the table, all panel members have wide knowledge in the field of research in general. The list includes experts in the systematic review, experts in the field of performance measurement and experts in public management. The ability to coordinate and communicate with all members of the panel was a challenge given the distant location of each from the location of the researcher who lives in another country. Gaining help from my supervisor to facilitate interaction with panel members was therefore essential.

Table 3 Review Panel

Name	Organization	Involvement
Prof. Michael Bourne	Cranfield University, SOM	Supervisor
Prof. Susan Vinnicombe	Cranfield University, SOM	Panel chair and expert in public management
Dr Jutta Tobias	Cranfield University, SOM	Panel member and expert in performance management
Dr Marek Szwejczewski	Cranfield University, SOM	Panel member and systematic review specialist

2.5.3 Identifying and Selecting Studies

This stage was concerned with identifying relevant articles that could provide answers to the review questions. This was based on the systematic review architecture provided by Cranfield School of Management. The following sections will provide an explanation about the first stage of the process which involves several steps comprising: (1) identifying initial sources; (2) generating a keyword list; (3) selection of databases; and (4) developing an assessment and quality criteria.

2.5.4 Initial Sources

A number of articles that were thoroughly studied and examined in previous stages such as critical review, mapping, and scoping study were used as a base for identifying relevant sources in the two main fields of performance measurement and public management. I started by identifying a list of papers- core and periphery- that seemed relevant to my field of interest and corresponded to the mapping process which represents the intersection between the two domains. A data extraction form was created and included general information about papers such as title, authors, date of publish and journal name. It also captured essential information based on a brief description about questions or problems raised, relevant theories and concepts, methods used, and papers' contribution to knowledge. This kind of information was the base for developing the summary list of papers that will be discussed down in another section. Accordingly, I was able to identify key relevant academic journals that publish relevant papers to my field of interest. Table (4) provides a list of journals that seemed relevant at the beginning of the systematic review.

Table 4 Initial Journal Selection

Publication Name	Ranking
Public Administration Review	4
International Journal of Operations and Production Management	3
Public Administration	3
International Journal of Public Sector Management	1
Management Decision	1
Measuring Business Excellence	NR

2.5.5 Databases

Online databases are a rich source of information that contains the most relevant, up to date and rigorous papers from very reliable and highly ranked journals. ABI/Proquest and EBSCO were selected as the primary databases suitable for the purpose of my research. The selection of these two databases was influenced by my reading from

previous stage of the scoping study and consultation with my supervisor. Several scholars such as Franco-Santos et al. (2007) and Franco-Santos, Lucianetti and Bourne (2012), have used and limited their search on these databases as they cover a large number of key journals and up to date articles in the management field which can hardly be compared to any other databases. Table (5) provides description about the two databases.

Database	Description*
ABI/INFORM(Proquest)	ProQuest Research Library™ provides one-stop access to a wide range of popular academic subjects. The database includes more than 5,060 titles—over 3,600 in full text—from 1971 forward. It features a highly-respected, diversified mix of scholarly journals, trade publications, magazines, and newspapers. This combination of general reference volume and scope makes it one of the broadest, most inclusive general reference databases ProQuest has to offer.
EBSCO	EBSCO serves the content needs of all researchers whether they access EBSCO resources via academic institutions, schools, public libraries, hospitals and medical institutions, corporations, associations, government institutions, etc. EBSCO offers more than 375 full-text and secondary research databases and over 515,000 e-books plus subscription management services for 360,000 e-journals, e-journal packages and print journals.
* Description taken from the databases' websites	

Table 5 The Search Databases

2.5.6 Search String

The selection of the search string was initially developed during stage 1 of the scoping study and the subsequent protocol review. It evolved over time and was based on several search attempts and reflection of results. Given the large number of published work in the field of performance measurement and management, the selection of key words was based on a careful attempt to narrow the result to a manageable number of papers suitable for investigation without losing relevant articles. Table (6) provides an indication about some of the search strings that were used and subsequent results from

the two main databases. It also provides some useful information about the size of a large domain such as performance management.

Table 6 Initial Search Strings

Initial Search Strings	EBSCO	PRO-Quest	Note
AB Performance AND AB (measur* or manage*)	59933	65126	Searching abstract
AB Performance AND AB (measur* or manage*) AND AB (public or government*)	5826	8181	Searching abstract
KW Performance AND KW (measur* or manag*) AND KW (public or government*)	233	2176	Searching using key words
ab(Performance) AND ab((Measur* OR manag*)) AND ab((public OR government*)) NOT ab(business* OR sale OR industr* OR firm* OR compan*)	5729	5557	Searching abstract

My selection of the search strings was then influenced by previous works on the field of performance management that were published in leading and well known journals. For instance, (Franco-Santos et al., 2007) conducted a systematic review to identify the key characteristics of business performance measurement systems and decided to focus on the phrase “business performance measurement systems” as their unit of analysis. Whilst their focus fell into the domain of business or the private sector in general, which is not the interest of this study; it provided a helpful insight on possible terms that could be utilized for our purposes, specifically those that relate to the field of performance measurement. Similarly, Fryer, Antony and Ogden (2009) conducted a literature review using key words such as “performance management”, “performance measurement” and “public sector”. Applying the same approach, that is using brackets to identify phrases such as “performance measurement” or “performance management”, could’ve been helpful in narrowing down the overall results, but since the numbers of retrieved papers were not sufficient I decided not to use brackets in my search. Furthermore, one of the reasons for using the word ‘management’ in addition to ‘measurement’ is the fact that many papers tend to mix between the two terms (Sole, 2009; Fryer, Antony and Ogden, 2009; Goh, 2012). Similarly, the selection of some other words such as the word ‘government’ in addition to the word ‘public’ was based on the same rational.

The field of performance measurement systems is a multi-disciplinary field (Franco-Santos et al., 2007) which might explain the large number of articles generated by initial searches. Given the purpose of the review questions additional words like ‘implementation’ and ‘adoption’ were then used to narrow the result further. Generic words such as ‘design’ and ‘use’ that are used in almost all papers were abandoned from the search process. Further, as can be seen from Table (7) the truncation character ‘*’ was employed to find words that might contain the same roots. For instance, a search for ‘measur*’ might yield different results like ‘measurement’, ‘measuring’, ‘measures’, and ‘measure’.

Table 7 Final Search String

Keyword	Rational
Performance	Central topic
And	
Measur* OR manag*	Main area of interest
And	
Public OR Government*	The context of the field of study
And	
Implement* OR Adopt* OR Appl* OR Driver*	The inner focus
Not	
Business* OR industr* OR firm* OR compan* OR Corporat*	Excluding papers about the private sector

2.5.7 Assessment and Quality Criteria

The search began with certain criteria to assess the value of results. Inclusion-exclusion criteria were necessary for setting boundaries and focusing the search in a way that corresponded to the purpose of the review. Many of the filters used were derived from previous examination of the literature during stage 1. For instance, the decision to take into consideration published papers following the year 1992 was influenced by the historical threads of performance measurement (Holzer and Kloby 2005; Folz,

Abdelrazek and Chung, 2009) and the work of Franco-Santos et al. (2007) and Franco-Santos, Lucianetti and Bourne (2012). Authors used 1992 as a ‘cut-off point’ because of the change in perspective in the field of performance measurement in that year, particularly following the introduction of the Balanced Score Cards (BSC) framework by Norton and Kaplan of Harvard University (Folz, Abdelrazek and Chung, 2009). Additionally, in order to give more value to the outcome of the review, the search was limited to papers that were published in peer-reviewed and scholarly journals only. Moreover, for obvious reasons that took into consideration the fundamental differences between public and private sectors as outlined in the scoping study, the search was also limited to papers that tackle the first rather than the latter. However, papers that did not make a clear distinction between the two sectors were included and considered relevant for the purpose of the review.

As can be seen in Table (8), papers were assessed based on their relevance on a scale of four points. Given the relatively large number of articles that tackle the field of PMS- even within the context of public organizations- and based on the review questions, both weak and completely irrelevant papers were ignored and delisted. Relevance was determined by papers’ ability to answer the review questions. Given the inevitable difficulty to find a satisfactory number of papers that cover all aspects of the field of interest as outlined by the review questions, even papers with small section(s) covering the topic were considered relevant and included in the review. Table (9) provides further explanation about filters that were used and the rational for each as stated in the protocol review.

Table 8 Papers' Relevance on a Four Points Scale

Relevance	Description	Value
Yes	the inclusion/ exclusion criteria are met	3
Somewhat	the inclusion/ exclusion criteria are partially met	2
Weak	the inclusion/ exclusion criteria are not met	1
Not relevant	the inclusion / exclusion criteria do not apply to retrieved paper	0

Table 9 Inclusion/ Exclusion Criteria

Criterion	Rationale for Inclusion
Research after 1990	The development of the current approaches of PMS
Academic publications on the topic of enablers or barriers of performance in public organizations	Relevance to the main topic
Publications with reference to design of PMS in public organizations	Relevant as it determine different designs and/or the experience of each organization in implementing these designs
Performance measurement vs. performance management	many papers tend to mix between the two terms
Criterion	Rationale for exclusion
All articles that speak about PMS in the private sector	The context is on public sector
The role and impact of performance measurement in public organizations	Relevance to the topic of interest
Theoretical and empirical assessment of the value and usefulness of performance measurement	Relevance to the topic of interest
Measuring the performance of human resources	The focus is on organizational performance and not on individual performance

Following the approach and recommendation of (Wallace and Wray, 2006) papers found were carefully assessed for their quality. A quality assessment grid was created upon which all papers were evaluated and was based on certain elements such as:

- Contribution: the degree to which these papers were adding original and meaningful contribution to knowledge and practice.
- Theoretical and Empirical Foundation: what are the explicit or implicit theoretical foundations? Do papers introduce empirical accounts for relevant frameworks and approaches in the field of interest?

- Methodology: the degree to which papers reviewed are able to reach a fit between research methodology and research question(s) (Tranfield, Denyer and Smart, 2003).
- Literature Review: Do these papers provide adequate literature review? Do they recognize and consider different opinions in the review?
- Ranking of Journals: ranking of journals was based on the eighth edition of Cranfield University Journal Recommendations for Academic Publication.

As stated earlier, applying certain filters and quality criteria deemed to be helpful in maintaining a balance between relevance and quality. However, I realized that applying such strict conditions might not yield a sufficient number of papers for my review. Through a discussion of findings with my supervisor and some panel members I was advised to consider looking into grey papers (i.e. papers that are yet to be published in academic journals) and papers from peer reviewed journals that are not necessarily part of Cranfield's list of ranked journals.

2.6 The Search Process

The search went through two main stages. The first stage was concerned with finding papers that are relevant to the research topic and pass the quality and assessment criteria. The outcomes of this stage represent the base for additional search in the following stage which was based on a citation and reference analysis for the purpose of finding additional relevant resources. The overall search process can be better explained through the systematic review architecture provided by Cranfield School of Management Figure (2). As can be seen from the figure the search process went through the following steps:

- The process started by conducting a comprehensive search on the main online databases using keywords that were developed from the pre-planned protocol review on the previous stage as explained above.
- Both online databases provide the possibility to focus the search on specific fields such as title, authors' name, abstract, given keywords etc. Based on consultation with my supervisor, a decision was made to focus the search on papers' abstract rather than papers' given keywords. That was mainly due to the high level of inconsistency between results from the two databases, where one database generated a large number of papers whilst the other generated very few. Table 10 provides an example about the results from an initial search on papers' supplied keywords using a simple search string.

Table 10 Example of a search using papers' keywords

Searching using supplied keywords	EBSCO	PRO-Quest
KW Performance AND KW (measur* or manag*) AND KW (public or government)	233	2176

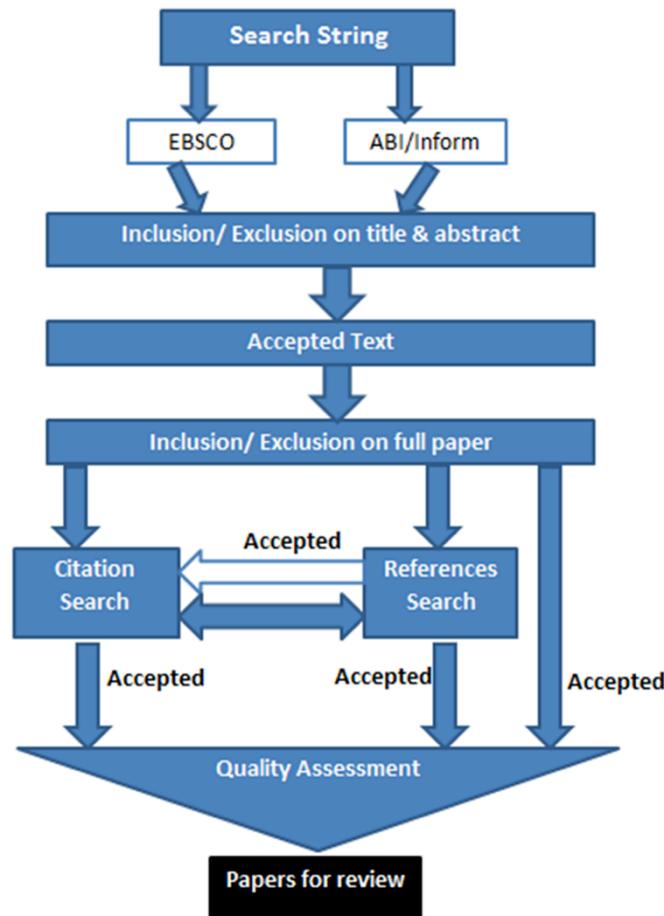


Figure 3 The Search Process

- The search results were narrowed down further by employing some of the search criteria, such as limiting my search to scholarly journals and setting a cut-off point of the year of publication which allowed the search to start from a specific year (1992), as explained in the assessment section. That was implemented through the use of special functions in the two main databases EBSCOhost and ABI/Inform (Proquest).

- Through the lens of pre-determined filters, retrieved papers were initially assessed and selected. That is to say, papers were screened by reading titles and abstracts only. Papers that met the inclusion criteria were considered fit for further scrutiny whereas those which did not were excluded from the review and considered irrelevant.
- Based on the former step, papers that passed the first level of scrutiny were collected for further assessment using an explicit selection criteria but this time through an analysis of the full content of each paper.
- Papers that were considered relevant would then proceed to a second round of evaluation to determine their quality. The basic details and reasons for exclusion were documented for all papers that were considered irrelevant, whether this was due to their quality or relevance to the research topic.
- I used special software for that purpose and for grouping relevant papers as well as for tagging each article with relevant remarks based on my reading. That proved to be helpful in latter stages as I needed to go back to some papers that met all criteria except for the journal ranking as explained in the criteria section.
- Two main data extraction forms were created- basic and detailed. While a basic data extraction form will capture essential information of each paper such as title, author(s), journal, and date of publication, the detailed extraction form included additional important information about papers such as the purpose of each paper, methodology, main findings, limitations, the contribution to the body of knowledge, and the degree of relevance to my topic.
- Papers that were selected for review were imported into NVivo software for analysis and synthesis. Those papers were also used to conduct the next round of search that is based on citations and references.
- For each paper, a citation search was conducted to extract all the articles that cited the same paper. Similarly, I conducted a reference search in order to identify all articles that were mentioned as a reference in all papers. Both citation and reference searches were conducted using the Web of Knowledge database. And, in the case where the Web of Knowledge was not able to generate result for a specific paper, the citation search was conducted using Google Scholar while references were extracted directly from papers.
- Retrieved articles then went through the same process of screening using the previous inclusion/exclusion criteria to identify additional suitable papers for review.
- The same process of screening and citation and reference searches continued until I reached a level of saturation and no more papers emerged from the search process as described by Figure (3).
- Finally, as explained above, all texts that emerged as relevant were imported to NVivo software for analysis and synthesis.

2.7 Final Result

A total of 837 papers were the initial result of the search process from the two main databases. The selection process went into two main stages based on relevance to the main questions and on the predetermined inclusion and exclusion criteria. The first stage of selection focused on evaluating papers based on title and abstract. A total of 91 papers resulted from the first level of screening. Those papers were then cautiously examined by a careful and critical reading of the full content. The following table shows that the majority of papers come from the ABIPROQUEST database.

Table 11 Search Results

Database	Reviewed	Selection based on title and abstract	Selection based on full text	Percentage (%)
ABIPROQUEST	503	59	33	6%
EBSCO	334	32	19	5%
Total	837	91	52	6%

By eliminating 25 redundant papers from a sum of 52 papers that passed the first and second level of screening the overall result set out to be 27 papers Table 12.

Table 12 Removing Redundant Papers

Summary Information	Total
Total Results	52
Duplicates	25
Final Results	27
Percentage	51%

Table 13 shows the additional sources that were included in the review. 10 sources were considered relevant from the previous stage, the scoping study. An additional 14 sources were included as the result of citation and reference analysis, 1 paper recommended by my supervisor, and finally 2 books giving a total of 57 sources to be included in the review. It is worth reiterating that the same criteria that were applied at the early stage of the search process were also applied on all sources all along.

Table 13 SR sources

Source	Source Details	Number
Articles	Database Search	27
	Scoping Study Papers	10
	Citations and references' search	16
	Recommendations	1
Books		2
PhD/Master thesis		0
Websites		0
Conference papers		1
Total Sources		57

2.8 Data Analysis and Synthesis

The final step in the process of systematic review is data analysis and synthesis Figure (1). Relevant papers were imported to special software for data analysis and synthesis. NVivo is a computer program for data analysis normally of qualitative type. The software provides useful means for organizing documents, coding data, and restructuring themes and thoughts. It also has other useful functionalities such as linking various documents, codes, and ideas. According to Hart (1998), while data analysis can be described in terms of breaking down something into its constituent parts and how these parts relate to each other, data synthesis is the act of making connection between parts identified in analysis in order to come up with a new arrangement or construct that has not been produced previously.

The first step in this stage was data analysis. The purpose was to investigate the literature and to extract ideas, theories, arguments, and methodologies for further synthesis. Sources were grouped into three main categories corresponding to the three sub-questions of the main question review: characteristics of successful design; drivers

for PMS implementation; and factors that affect the adoption and implementation of PMS. The aim was initially to consolidate my understanding of each construct before linking these categories together. The ability to identify similar or new themes as well as grouping them into certain categories was facilitated by the software which allows an instant coding of materials. So, bearing in mind the review questions, coding begins as soon as certain ideas or themes emerges in the text. It was possible to implement this instantly while reading the text through a special function in NVivo. There was also the possibility to create categories and subcategories based on the findings from papers. Common themes from the same study or from other texts were grouped into a certain coding structure that corresponded to the review questions. A continuous refinement of all themes and sub-themes was conducted in relation to findings from the review materials.

2.9 Descriptive and Thematic Analysis

Essential information from each paper were collected in a data extraction forms pertaining to the main systematic review question and sub-questions as per Appendix 1. A detailed extraction form was also created for all papers which include the pre-mentioned information in addition to useful summary of all papers. While the data extraction form provides useful but general information about papers the detailed content forms are concerned with the contribution of each paper. Main arguments and issues from each paper are registered for further analysis. Comments about each paper findings and methodology were also part of the detailed extraction form. These forms were helpful for a later stage of data analysis and synthesis. This section provides further detail about the findings of the systematic search process.

2.9.1 Descriptive Findings

Publication Features

As can be noted from the figure (4) the systematic review was mainly dependent on papers from scholarly and peer reviewed journals.

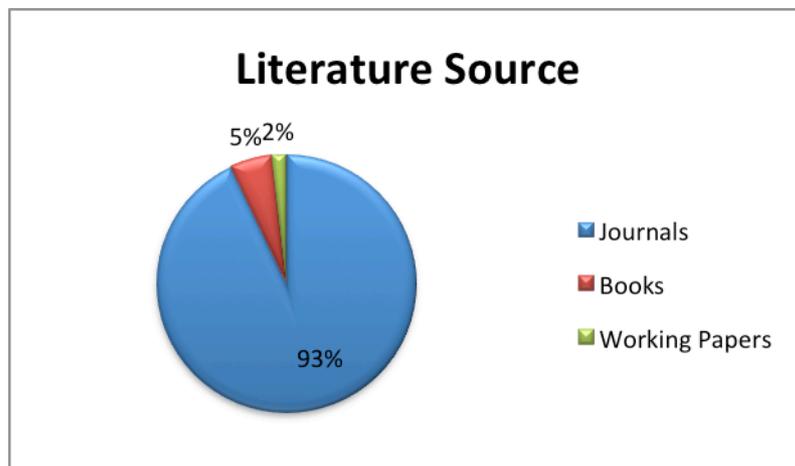


Figure 4 Literature Source

Used Journals

The 51 papers that were included in the review were obtained from the following journals in Table 14. The table shows the number of papers that were retrieved from each journal and their ranking according to Cranfield's Journal Recommendations.

Table 14 Journals and ranking

Publication Name	No. Papers	Ranking
Accounting, Organizations and Society	1	4
Public Administration Review	7	4
Accounting Horizons	1	3
International Journal of Operations and Production Management	1	3

Local Government Studies	2	3
Public Administration	2	3
International Review Of Administrative Sciences	1	2
Journal of Public Administration Research and Theory	3	2
Nonprofit Management and Leadership	1	2
Production Planning and Control	1	2
public administration and development	1	2
International Journal of Public Sector Management	10	1
Management Decision	1	1
Public Performance and Management Review	2	1
Systemic Practice and Action Research	1	1
Total Quality Management and Business Excellence	1	1
Asia-Pacific Management Accounting	1	NR*
Australian Journal of Public Administration	1	NR
Evaluation and program planning	1	NR
Financial Accountability and Management	1	NR
International Journal of Health Care Quality Assurance	1	NR
International Journal of Productivity and Performance Management	2	NR
International Review of Public Administration	1	NR
Journal of Accounting and Organizational Change	1	NR
Journal of Healthcare Management	1	NR

Measuring Business Excellence	4	NR
Procedia - Social and Behavioral Sciences	1	NR
Total Number of Papers	51	
*NR not ranked		

While the majority of articles are from ranked journals around 31% are from unranked journals Figure (5). As indicated in a previous section, the decision to include unranked journals came after consultation with my supervisor and some members of the panel and was merely due to the insufficiency on the number of papers needed for the review. However, for quality purposes, only scholarly and peer reviewed papers were included in the review.

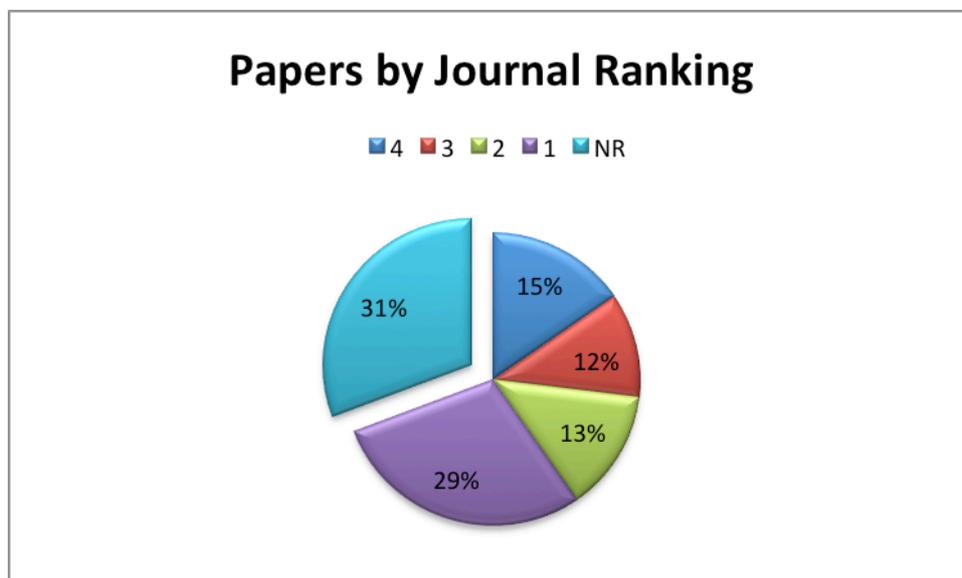


Figure 5 Papers by journal ranking

Applied Framework

Papers reviewed show that the BSC is a dominating framework that has exhibited significant growth over the past decades. 13 papers out of the 20 papers that mention PMS type were about BSC. Figure (6) shows the percentage of papers that mention other PMS designs compared to those that mention BSC.

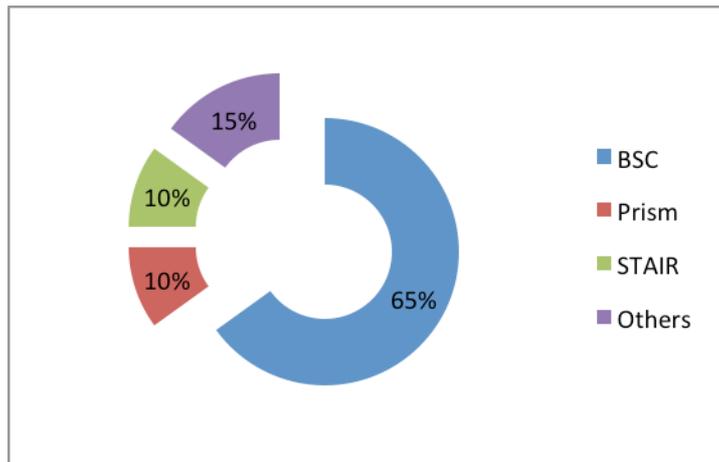


Figure 6 PMS design

Field of interest

Findings from the search process indicate that the majority of papers reviewed were concerned with the determinants of actual development and use of measures in public organizations Figure 8. Few papers were concerned with the motivation behind adopting PMS in the public domain. One possible explanation for the limited focus on this topic is that the rationale for measurement is rarely questioned and consequently the supply of performance information is taken for granted (Van Dooren, 2005). Similarly, limited numbers of papers were concerned with explicitly identifying the characteristics of a successful measurement design. These characteristics are implicitly assumed rather than explicitly discussed. That might be related to the limited empirical studies in the field of PMS within the context of the public sector and the limited number of frameworks that are designed solely for public organizations.

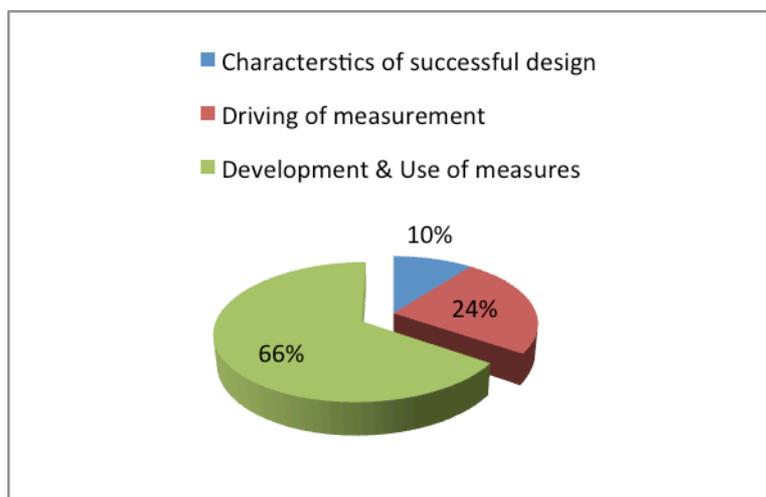


Figure 7 Papers by the field of interest

Papers' Methodology

The number of research papers that are based on qualitative or quantitative techniques can be shown in Figure (8). While findings suggest that empirical research is based on a balanced use between quantitative and qualitative research strategy, it clearly shows that a mixed approach is rarely used. A qualitative approach was normally used to grasp and explain a particular phenomenon, and the quantitative approach in reviewed papers was used in developing and testing particular hypotheses related to factors that affect the utilization of measurement. Additionally, Figure (9) allows for a further examination by looking into the timeline of research in terms of the use of qualitative, quantitative, or mixed methods. The figure could indicate that utilization of PMS in public organizations is still an emergent issue since the focus continues to shift between exploratory and theory testing methods.

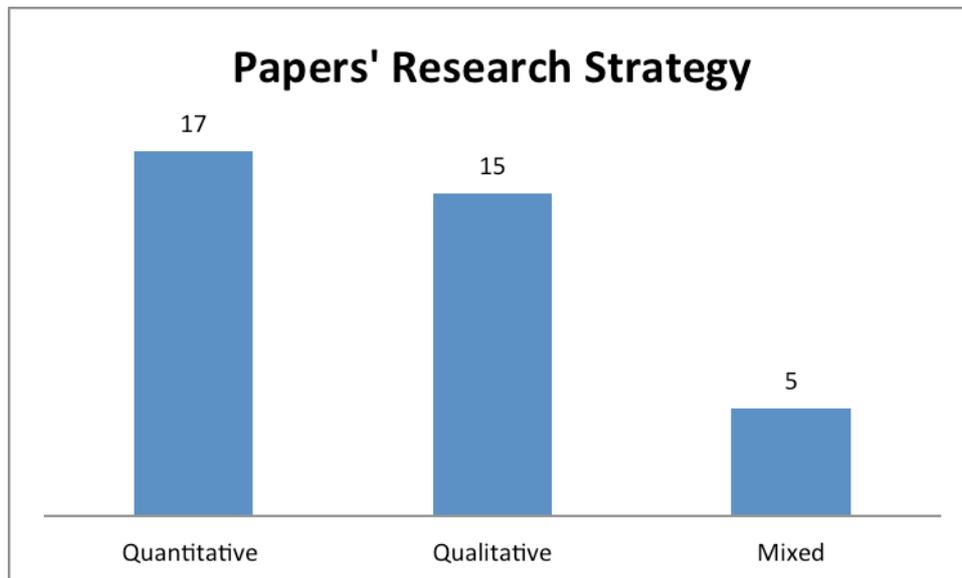


Figure 8 Research strategy

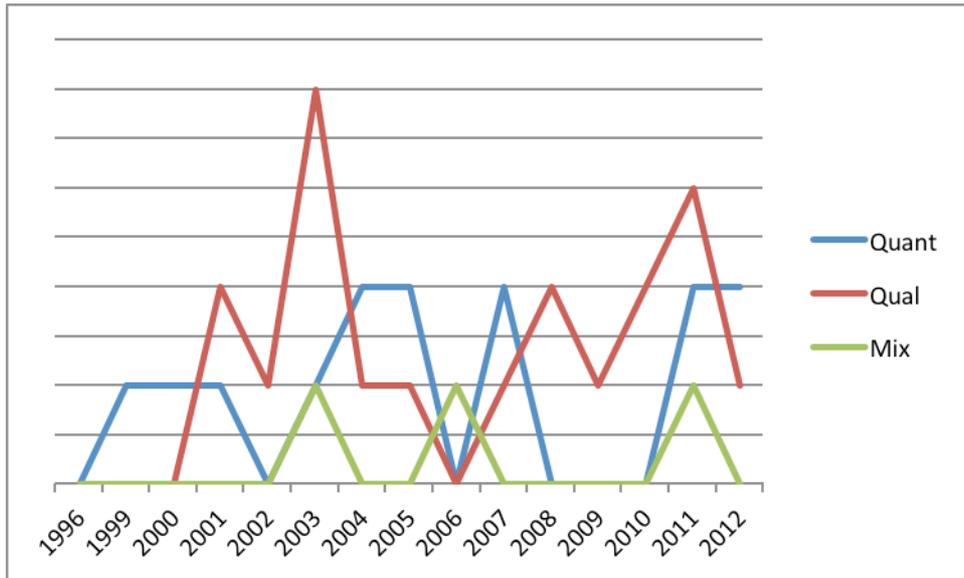


Figure 9 Research strategy timeline

Publication Year

Comparing sources by the year of publication, Figure (10) shows a growing interest on the field of performance measurement and management in the public domain over the past two decades. As can be seen from the graph, fewer papers that are relevant to the field of interest were published prior the year 2000. The increased number of relevant papers in the past few years is an indication of the proliferation of performance measurement and management in public organizations and, as a consequence, an increased interest in implementing change and measuring its impact.

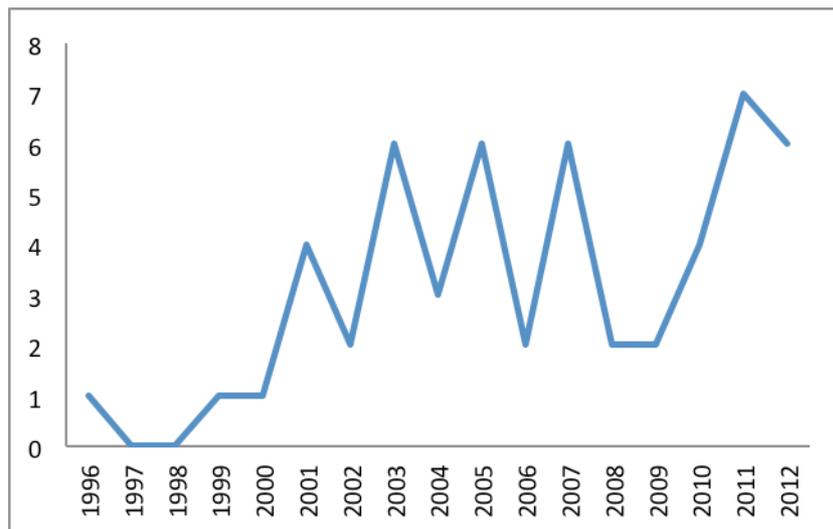


Figure 10 Papers by publication year

Further, Figure (11) shows a shift in the interest of academic literature from the initial focus on performance measurement into a more general concept of performance management. The shift could be attributed to a variety of reasons such as the maturity of measurement use in private and public organizations and to the evolution of measurement systems such as the Balanced Scorecard of Kaplan and Norton 1992. *“Originally, we thought the Balanced Scorecard was about performance measurement..., we quickly learned that measurement has consequence far beyond reporting on the past... Balanced Scorecard concept evolved from a performance measurement system to become the organizing framework, the operating system, for a new strategic management system”*(Kaplan and Norton, 2001. P99).

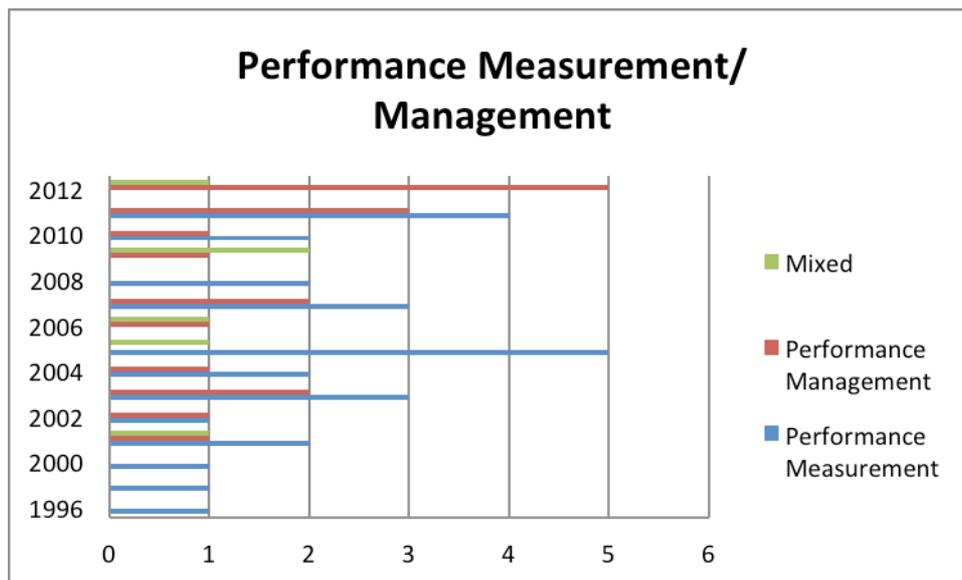


Figure 11 Performance measurement vs. performance management

Type of Knowledge

Papers were categorized by knowledge-type, using Wallace and Wray (2006) classification. According to authors, theoretical knowledge papers develop a systematic reflection of the social world, research papers depend on a systematic investigation of the social world, and practice knowledge involve the actions that might be taken in the social world. As can be seen from Figure (12) the review was influenced by theoretical and research papers.

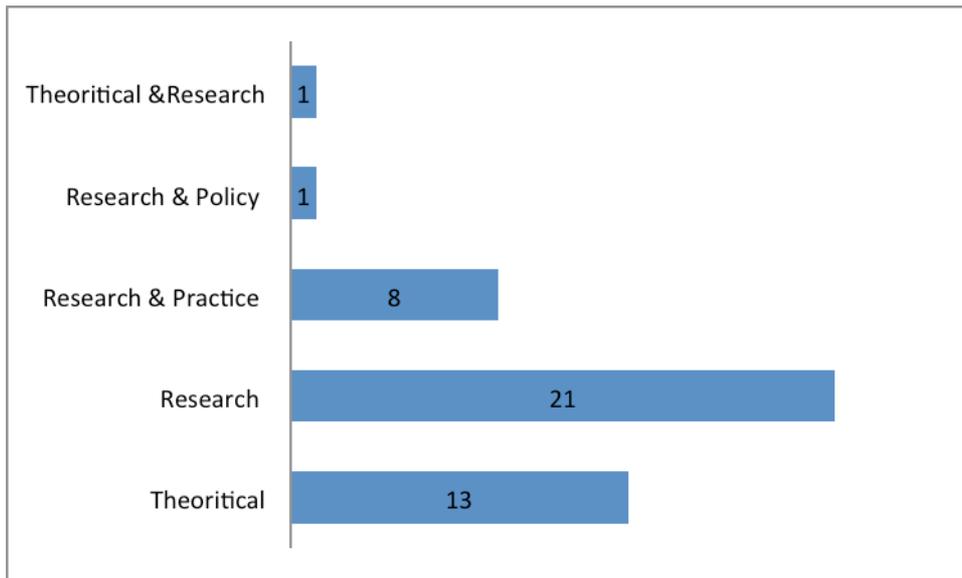


Figure 12 Papers by the type of knowledge

Sources by Country

As illustrated by Figure (13), retrieved research papers pertain to developed countries, mainly the USA and UK. The descriptive findings could therefore imply a cultural bias. Only one research empirically examined the implementation of PMS outside that context- a public organization in Croatia. This confirms the need for further research in other contexts, such as developing countries.

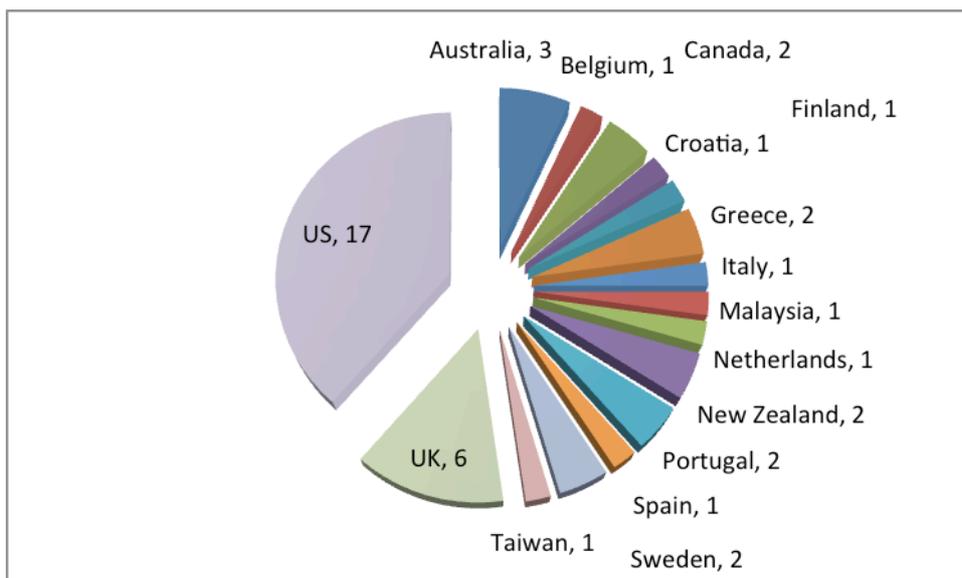


Figure 13 Sources by Country

Additionally, the figure can also indicate the active research community in these countries and the dominance of the USA and UK over the field of research.

2.9.2 Systematic Review- Limitations

The selection of relevant sources of information went through different stages and was not a straight forward process. The main limitations of the SR may be related to the following:

- The systematic review process might have a negative influence on the ability to fetch additional papers in relevant topic(s). Despite their relevance, the process cannot justify papers that might be acquired by chance, intuition, or previous random readings.
- The systematic review was mainly based on papers from scholarly journals while other sources such as books were disregarded at the early stage of the process. The decision to exclude books was because of concern of time, accessibility, cost and quality- given that a large number of books have been published in the field of PM while the quality of each is hard to determine.
- The time span of the systematic review process presented several challenges. The primary challenge is related to the need to continually update the list of relevant articles which had a consequent impact on the overall analysis and synthesis of data.
- It was also difficult to determine the level of saturation where no more relevant papers can be found from the process of citation analysis and cross referencing. After several attempts I was convinced that no more relevant articles will emerge from that process only to realize that additional sources continue to come out.
- The inconsistency between the two databases in terms of the number of papers that emerge based on a search using keywords Table (10) was an additional challenge. For it forced me to search in abstract which is not as accurate as authors' supplied keywords.

Inconsistency in terms of authors' different keywords was a major challenge. It created confusion and required revisiting the literature several times. The issue was resolved by a careful analysis of specific terminologies for all papers and by relying on my own interpretation and synthesis of the overall meaning.

3. Determinants of Successful Utilization of PMS

3.1 Introduction

The systematic review worked to identify determinants of successful utilization of performance measurement in public organizations by examining key features and factors pertaining to different phases in the utilization process. Topics that have been under examination and cover the area of performance measurement include the development, adoption, implementation, and use of measurement (Mimba, Helden and Tillema, 2007; Cavalluzzo and Ittner, 2004). However, the majority of academic papers do not explicitly define the meaning or the boundaries of these terminologies. For instance, 'implementation' seems to be a fuzzy word that could have different meanings. While it refers to the process of measures' development (Berman and Wang, 2000; Cavalluzzo and Ittner, 2004; Sanger, 2008; De Waal, 2003), it also refers to the actual use of measures for various purposes (Van Dooren, 2005; Julnes and Holzer, 2001; Moynihan, 2004; Yang and Hsieh, 2007). Similarly, adoption could be viewed in terms of the development of a capacity to act (Julnes and Holzer, 2001) or more broadly as the set of behaviors which influence the decision to apply measures and their development (Van Dooren, 2005). Further, many studies have been less clear about the distinction between adoption and implementation and there is no conclusive evidence regarding the determinants of both (Yang and Hsieh, 2007). This has created a confusion about the exact meaning of these terms (Mimba, Helden and Tillema, 2007) and adds to the difficulty of identifying factors pertaining to each level of the utilization process. It also shows the complexity of research in this field and makes it harder for researchers to generalize findings and draw conclusions based on careful analysis and comparisons between studies. Researchers need to be mindful of the different usages of these terms and their determinants in order to be able to elicit useful information.

Following the discussion I had with my panel, utilization of measurement in this paper was viewed in terms of two stages- adoption and use of measures. Adoption is about having performance measurement which starts by general recognition about value and benefits of a measurement initiative, and ends by the selection and development of measures. The second stage is related to the actual use of measures for making decisions in public organizations. However, following the content analysis it was noted that both the development and use of performance measures/indicators are affected by the same determinants. That is to say that both stages are affected by the same groups of factors, albeit with a varying degree of influence.

Bearing that in mind, and to avoid further confusion, this paper uses the supply and consumption analogy to identify determinants of successful utilization of performance measurement in public organizations as outlined by Van Dooren (2005). This implies a

distinction between the production process and uses of performance information (PI). Consequently, pursuing the first sub-question of the systematic review, adoption would be viewed in terms of: the driving factors for applying measurement (the demand side) and factors that affect the development of measures (the supply side). Additionally, whereas the development of measurement represents the supply side, actual use of measurement for decision-making represents the consumption side.

3.2 Adopting PMS in the Public Sector

Few papers in the review explicitly determined the driving factors for adopting PMS in the public domain. One possible explanation for the limited focus on this topic is related to the fact that performance measurement has become a pervasive feature and an important component of any reform program in the public sector (Torres, Pina and Yetano, 2011; Folz, Abdelrazek and Chung, 2009). Given its reputation to improve performance and its wide acceptance in private companies and due to the continuous importation of practices from private to public organizations, led by the school of New Public Management (NPM), the rationale for measurement is rarely questioned and subsequently the supply of performance information is taken for granted (Van Dooren, 2005). However, papers reviewed suggest that looking into these factors could be helpful for several reasons. First, as noted by Moynihan (2004), the apparent failure of adopters to actually use measures begs the question of why they were so widely adopted in the first place. Examining the driving factors for adopting measurement could be helpful in providing an answer to this question. Secondly, examining factors that affect management decision to develop performance measures or indicators should help public sector managers to better plan the application of a new measurement system in their organizations. It enables them to put better incentives and programs for the successful use of measures. This is based on the well-known assumption that the purpose of any system(s) or change program should determine the use of such system or program (Behn, 2003). Finally, identifying the motivation for adopting PMS is crucial not only to determining its use but also the level of use. For instance, one implication for the externally imposed PMS is that it is likely to be ineffective or less likely to influence internal use. Managers and politicians are usually reluctant to disclose performance information if it was originally designed for decision-making (Torres, Pina and Yetano, 2011).

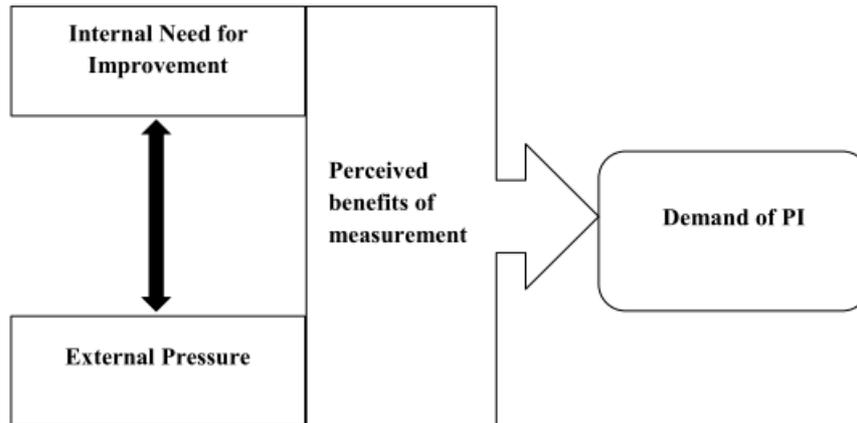


Figure 14 Demand of PI in public organizations

Papers in the review indicate that demand for performance information in public organizations is based on two main driving factors: an internal desire to make better decisions and improve performance; and pressure from external stakeholders. It shows that both internal and external political groups are motivated by the value they accomplish with measurement initiatives as shown in Figure (14).

3.2.1 Internal Need for Improvement

The demand for performance information is based on the need for making change and its influence on the overall performance. The systematic review suggests that, in a public organization, PMS is more likely to be adopted when the policy of using the system comes from within the organization as an internal requirement. It shows that an internal decision to apply PM is a rational approach that is based on a calculation process that works to determine the value of a new system application. An actual desire for improvement is normally driven by top and middle management based on their assessment and valuation of both current practice and the benefits of the new system (Moynihan, 2004). If internal actors realize the potential benefits of the measurement system they will be more inclined towards its application. Conversely, management will be less likely to adopt the new system if it is to think that it is going to undermine their power or negatively affect their interest. The consistency between rational decision for adoption and expectation has its direct influence on both the development and use of measures. It is an indication that chief executives do not consider measure as symbolic but with actual benefits (Folz, Abdelrazek and Chung, 2009).

Identifying the value of measurement is important in many ways. Firstly, understanding why a public manager measures performance is important to determine which measure

fits which purpose. For instance, measuring inputs and outputs could be useful to determine efficiency but less helpful to determine effectiveness which might be usefully measured by focusing on the outcome (De Bruijn, 2007; Behn, 2003). Secondly, the use of performance measures is determined by the way that public managers perceive their functions (Behn, 2003). Authors like Julnes and Holzer (2001) provide further explanation by indicating that the level of knowledge in the organization about the usefulness of performance measurement information is essential to assess an organization readiness to develop and use performance measures. Finally, it is worth noting that the perceived benefits are not only limited to the direct influence of the measurement function. An organization's interest in performance measurement could be related to the indirect benefits of adopting PMS. For instance, in the experience of many public organizations, performance measurement would require a clear formation of strategies, translation of all goals and objectives, and identification of measureable targets and results (Kaplan and Norton, 2001; Kaplan, 2001). Accordingly, and from a managerial point of view, these advantages that come from the application of measurement could be an end on their own right (Julnes and Holzer, 2001).

3.2.2 External political pressure

Pressure from elected officials, responding to citizen demands for greater accountability, and compliance with certain laws, administrative regulations, or reforms are some examples of external influences- in the form of demand for performance data- which according to many papers in the review will lead to the application of measures in public organizations. According to this view, the value of measurement and consequently the decision about adoption have been made externally with no or little participation from internal actors. Managers in these organizations have no choice but to comply with these requirements by developing and reporting performance to specific internal and external stakeholders. Their view along with the view of employees about the value and benefits of a new system of measurement is not, or slightly, considered.

The Modernizing Government Act (MGA) is a representation of an effort by the government of Greece to converge with the European common public management policies and regulations which call on to every public organization to set goals, measure performance, and report on results (Sotirakou and Zeppou, 2006). Other governments' initiatives with similar doctrines can be seen in different countries as noted by many papers such as the Government Performance and Results Act (GPRA) in the USA, the Local Government Act (LGA) in the UK, and the Public Service Act (PSA) in Australia. The main principle behind these initiatives is: adopting performance measurement will enhance efficiency and results and will improve responsiveness towards citizens (Cavalluzzo and Ittner, 2004).

This is perhaps why some papers choose to use the principles of institutional theories to explain adoption of measurement (Sotirakou and Zeppou, 2006; Cavalluzzo and Ittner, 2004; Taylor, 2011a; Torres, Pina and Yetano, 2011; Julnes and Holzer, 2001). The theory assumes that organizations adopt their structures and practices according to their

institutional environments. It claims that organizations' operations are guided by institutional criteria- not technical efficiency- in an effort to gain legitimacy (Sotirakou and Zeppou, 2006). According to those papers, adoption of performance measures can be viewed as a process of formal compliance with the wishes and expectations of the external environment and stakeholders and represents a sign of responsiveness and good management. However, if the focus in this scenario is not necessarily on the efficiency of the system but rather on the satisfaction of some external stakeholders, what would be the possible implication on the actual development and use of PMS in the public domain? Most papers in the review did not seem to be interested in providing answer to this question. Few others on the other hand warn that systems used to satisfy external requirements only are less likely to influence internal behavior (Cavalluzzo and Ittner, 2004). The issue will be discussed further in a following section.

3.3 Development and Use of Measures- PI Supply and Consumption

The factors selected from the literature are presented in Table (15). In reviewing the determinants of a successful implementation and use of performance measurement in public organizations, the initial analysis began with extracting relevant factors for the utilization process. The aim was to identify the main conditions needed for an effective development and use of performance measurement systems. Following a careful reflection and analysis of papers and various factors that were collected, a couple of issues must be noted. First, most papers in the review were concerned with elements that are more relevant to their particular organization(s) or context. Hence, different papers had different perspectives in terms of elements that determine success. Factors that were extracted from the literature demonstrate the lack of consensus on specific class or categorization of these factors. Second, based on my analysis there were no significance differences between different levels of utilization of measurement in terms of the main groups of factors that determine the success of measurement initiatives. That doesn't necessarily mean that similar factors will have similar influence over the development and use of measures. Rather, it shows that utilization of measures is affected by the same groups of factors but with varying impact on the supply and consumption of performance information. It could be that the significance of each group of these factors are determined by other elements notably the maturity of an organization in relation to the use of measurement and the level of understanding of the measurement value and function.

The general review of the literature indicates that utilization of PMS in public organizations is determined by three main groups of factors: political, technical, and cultural Figure (15).

Political factors are related to the role of stakeholders that normally exercise some control and pressure to achieve certain aims. It encompasses the influences from inside and outside a public organization. While internal interest groups are mainly management and employees, external interest groups on the other hand might include but not limited to elected officials, regulators, labor unions, and citizens. A clear majority of the reviewed papers is devoted to highlight the importance of political actors and their role to determine the success or failure of measurement initiatives. However, while understanding that political context is necessary for effective adoption and use of measurement (Julnes and Holzer, 2001), the multiplicity of internal and external players could add many challenges in public organizations.

Technical factors are related to the available capacity for an organization to develop and use PM. It includes the adequacy of available human and financial resources. Technical assistance, training, and mentorship are important elements of this category. Issues like the quality and appropriateness of the measurement design were considered relevant technical factors as well. Further, these factors include elements of uncertainty in terms

of defining and developing measures due to the complexity of operations and functions in the public sector.

Cultural factors encompass all the elements that define and shape an organization's culture which has an impact on the development and use of measurement. Examining these cultural aspects was not easy since it involves the set of behaviors, values and principles inherent in the system, including operational, social, and institutional culture (Hawke, 2012). Alternatively, the focus was on those elements that shape an organization's culture which are considered a necessary base for effective adoption and use of measures.

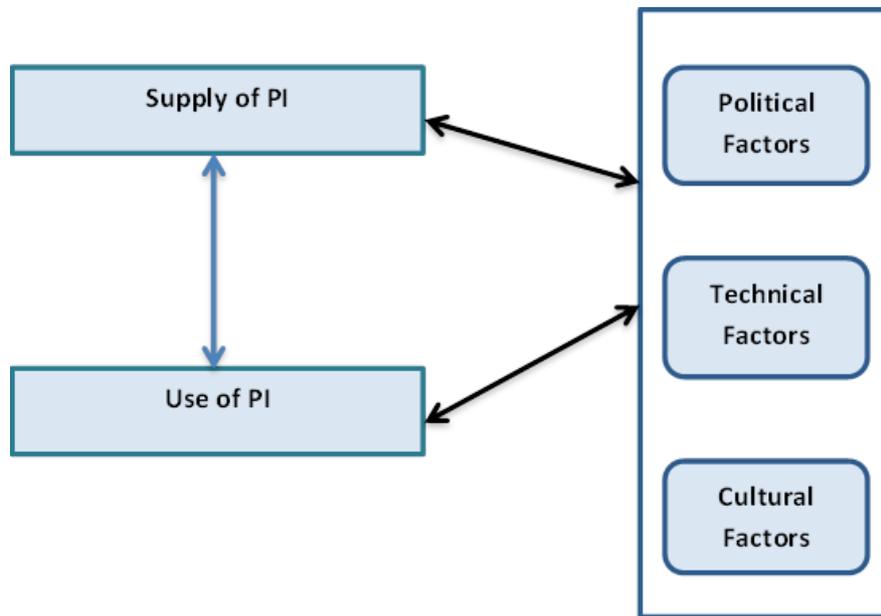


Figure 15 PI Supply and Consumption

Bearing in mind the impact of each individual factor on the supply and use of measurement, it is important to note the link that might exist between factors. These factors might interact to influence the development and use of measures (Cavalluzzo and Ittner, 2004). One or more factors or sub-factors can alter or accentuate the effect of others (Hawke, 2012). For instance, culture defines the way in which organizations outline and achieve their goals, but at the same time is often shaped and defined by both internal and external political agents. Similarly, while technical factors are dependent on the capacity and support of political interest groups they also affect the behavior of people and their perception about the value and benefits of change. However, most papers in the review did not provide a tentative discussion about this issue. Given the large number of potential interactions and limited theory on how these factors interrelate (Cavalluzzo and Ittner, 2004) there is clear deficiency in our knowledge in terms of the significance of these factors. It is difficult to identify generic factors from those that are

limited to few particular organizations and subsequently it is difficult to generalize findings. Additional studies should be mindful of the interrelationships that might exist between factors and the significance of their interaction over the utilization of measurement. Further, the utilization of PMS is not a singular event which can be executed automatically. According to Van Dooren (2005), it is not common or even expected to apply performance measurement at all levels and functions of a public entity all at the same time. Author explains that both adoption and use of measures are related and more of the first usually implies more of the latter and vice versa. Hence, identifying factors that affect the demand and supply of measures should work as a guideline for better consumption. And similarly identifying factors that affect consumption could explain the growing demand and supply of performance information. Public organizations out to perform a self-assessment against such factors and work purposefully towards better utilization of the system (Julnes and Holzer, 2001).

Table 15 Development and Use of measures

	(Adcroft and Willis, 2005)	(Berman and Wang, 2000)	(Bianchi and Rivenbark,	(Carmona and Grönlund,	(Cavalluzzo and Ittner, 2004)	(Lilian Chan, 2004)	(DeGroff et al., 2010)	Moynihan, Pandey and Wright (2011)	(Van Dooren, 2005)
Political Factors									
Leadership support and internal management commitment		x	x		x	x		x	
Employee buy-in and involvement		x				x			
External pressure		x			x				
Technical Factors									
Resources issues (including financial and human resources, time, training and technical knowledge)		x			x	x			x
Design issues (refers to the suitability of the measurement system)		x			x	x			
Measurement Issues (relates to the ability to measure performance in public organizations)	x			x			x		x
Cultural Factors									
Performance oriented Culture (including features such as clarity of goals, focus on end results, and the use of rewards and incentives)					x	x		x	x

Openness to change (refers to organizational orientation towards learning and development and the level of resistance)										
Other cultural considerations (mainly in non-western countries such as low institutional capacity, corruption, informality)										

<i>Continue Table 15</i>	(Folz, Abdelrazek and Chung, 2009)	(Goh, 2012)	(Hawke, 2012)	(Ho, 2005)	(Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006)	(Inamdar, Kaplan and Reynolds, 2002)	(Julnes and Holzer, 2001)	(Kwon and Jang, 2011)	(Lee and CHO, 2011)	(Mansor et al., 2012)
Political Factors										
Leadership support and internal management commitment	x	x	x	x		x	x		x	x
Employee buy-in and involvement	x			x		x	x			x
External pressure		x	x	x			x			
Technical Factors										
Resources issues (including financial and human resources, time, training and technical knowledge)			x		x	x	x		x	
Design issues (refers to the suitability of the measurement system)			x			x	x			

Measurement Issues (relates to the ability to measure performance in public organizations)									X	
Cultural Factors										
Performance oriented Culture (including features such as clarity of goals, focus on end results, and the use of rewards and incentives)			X			X	X	X	X	X
Openness to change (refers to organizational orientation towards learning and development and the level of resistance)		X								
Other cultural considerations (mainly in non-western countries such as low institutional capacity,		X	X		X		X			

corruption, informality)										
-----------------------------	--	--	--	--	--	--	--	--	--	--

<i>Continue Table 15</i>	(Mimba, Helden and Tillema, 2007)	(Moynihan and Pandey 2010)	(Moynihan, 2004)	(Northcott and Taulapapa, 2012)	(Pereira and Melão, 2012)	(Perera, Schoch and Sabaratnam, 2005)	(Radnor and Lovell, 2003)	(Rantanen et al., 2007)	(Rodgers and Keynes, 2011)
Political Factors									
Political Support (Leadership support and internal management commitment)		x	x	x	x				
Employee buy-in and involvement				x	x				x
External pressure		x	x				x		
Technical Factors									
Resources issues (including financial and human resources, time, training and technical knowledge)	x	x	x	x	x	x		x	x
Design issues (refers to the suitability of the measurement system)					x		x	x	x

Measurement Issues (relates to the ability to measure performance in public organizations)						x				
Cultural Factors										
Performance oriented Culture (including features such as clarity of goals, focus on end results, and the use of rewards and incentives)		x					x		x	
Openness to change (refers to organizational orientation towards learning and development and the level of resistance)				x						x
Other cultural considerations (mainly in non-western countries such as low institutional	x				x			x		

capacity, corruption, informality)										
--	--	--	--	--	--	--	--	--	--	--

<i>Continue Table 15</i>												
	(Sanger, 2008)	(Šević, 2005)	(Sharma and Gadenne, 2011)	(Sole and Schiuma, 2010)	(Sole, 2009)	(Sotirakou and Zeppou, 2010)	(Taylor, 2011a)	(Taylor, 2011b)	(Tillema, Mimba and Verheijen and 2005)	(Yang and Hsieh, 2007)	Total Percentage	
Political Factors												
Leadership support and internal management commitment)	x		x		x	x	x			x	x	60
Employee buy-in and involvement				x	x	x				x	x	37.5
External pressure	x										x	27
Technical Factors												
Resources issues (including financial and human resources, time, training and technical knowledge)	x			x	x	x	x				x	60
Design issues (refers to the suitability of the measurement system)			x			x	x	x				35
Measurement Issues (relates to the ability to measure performance in public organizations)		x					x					20
Cultural Factors												
Performance oriented culture (including features such as clarity of goals, focus on end results, and the use of rewards and incentives)		x	x	x	x	x		x				47.5

Openness to change (refers to organizational orientation towards learning and development and the level of resistance)		x				x	x					15
Other cultural considerations (mainly in non-western countries such as low institutional capacity, corruption, informality)	x								x			22.5

3.3.1 Political Factors

The reviewed literature indicates that successful implementation and use of performance measures in public organizations should take into account the impact and influence of different political groups (stakeholders) whether inside or outside an organization. It suggests that stakeholders' support and involvement is fundamental in order to achieve that end. It is central to the establishment of agreement on the system's purpose and key goals and objectives (DeGroff et al., 2010). If major stakeholders have no voice in performance measurement, they will be less likely to support it (Yang and Hsieh, 2007). However, 'stakeholders' is a broad term that might encompass all people that are part of the process of performance measurement (Goh, 2012). For instance, it might be applicable to those who: develop measures, manage the process of measurement, collect and distribute information, and benefit from and use measurement data. It also might be applicable to many others who may be affected by measurement indirectly such as citizens.

Bearing in mind the complexity of such definition and the difficulty of identifying all the stakeholders of a public organization, the majority of papers on the review particularly concentrate on the impact of internal political actors. Internal interest groups share different responsibilities such as designing the measurement system, setting performance indicators and targets, collecting and analyzing data, reporting performance results, and ultimately taking decisions' based on performance information (Cavalluzzo and Ittner, 2004; Folz, Abdelrazek and Chung, 2009; Moynihan, Pandey and Wright, 2011; Sanger, 2008; Yang and Hsieh, 2007). Most studies focused on the role of leadership and internal management commitment as well as the role of employees. External politics on the other hand were mentioned in terms of politicians, citizens, labor unions, funding agencies, legislators, and, more broadly, governments. The discussion in the following section is presented on the basis of an analysis that was carried out to distinguish internal and external powers in terms of their impact on the development and use of measures.

Internal politics

The literature indicates that utilization of PMS will not be successful in public organizations without proper support and commitment by internal interest groups mainly managers and employees. As can be noted from table 13, managerial support and leadership is a prerequisite for both adoption and use of measurement. An effective introduction and application of performance measurement requires committed leaders willing to provide significant

managerial investment and support (Sanger, 2008; Goh, 2012; Hawke, 2012). Cavalluzzo and Ittner (2004) point to the role of top management support in creating a suitable environment for change. According to them, top management commitment is expected to influence both employees' accountability to results and their use of information for decision-making. Authors also indicate that employees are more likely to consider change in the system favorable if they perceive strong support by top management. The majority of papers indicate that lack of managerial time and attention will not only contribute to the unsuccessful use of a measurement system but equally important impede its adoption (Lilian Chan, 2004). This should not be surprising, given the leverage that leaders have in their organizations and the outcome of their decisions. Senior managers are expected to conduct several vital functions that are essential for the success of any reform or change program. Their responsibilities will include but not limited to assigning and managing roles and functions, mobilizing resources, focusing on result and value, encouraging learning, empowering employees, and ultimately rewarding and sanctioning good and bad performance (Sanger, 2008). It should be noted that the support of managers or leaders in an organization is needed prior to and after the implementation of change. This gives an indication of the interrelation between internal politics and other technical and cultural considerations. For instance, internal management plays a key role in managing the planning process, creating the information infrastructure, selecting measures, and evaluating the results (Bianchi and Rivenbark, 2012). It is to that end why it was not possible for the majority of papers in the review to neglect their political influence.

Non-management employees matter as well. The application of a new system will impact them directly and is more likely to affect their interest. The introduction of a new system such as PMS will entail a major shift from normal practice and will require a lot of time and effort. Members of an organization might feel insecure about change, especially if they can't see its value or if they think that they will lose their interests. In order for performance measures to be implemented successfully, internal employees need to be receptive rather than being fearful (Folz, Abdelrazek and Chung, 2009). A key component to ensure the support of employees is by involving them in the development phase of measurement. Papers in the review suggest that a participative pre-implementation decision process is an efficient way to ensure the support of employees and the success of performance measures (Northcott and Taulapapa, 2012). It shows that their participation in internal political activities aimed at promoting measurement has a positive influence on their adoption (Julnes and Holzer, 2001). Involving employees in the development process will not only secure their support but also influence their views about PMS as a tool for improvement and learning. It creates ownership and enhance communication about targets and measurement processes (Goh, 2012). It also boosts motivation and encourages employees' accountability, ability to learn, and contribution to

solution (Sole, 2009). However, while some papers like Kaplan (2001) and Torres, Pina and Yetano (2011) stress on the bottom approach to measurement development many of the reviewed papers didn't explain how involvement might be achieved. The literature didn't offer enough discussion about the way to involve employees and how to manage their involvement. One concern is that the large pool of participants could add to the complexity of managing the views and demands of large number of people inside an organization. Involvement might turn to be a burden rather than facilitating implementation.

Very few papers were keen to provide answers to questions like: what would determine the support of internal interests groups and how to ensure that an organizations' leadership will continue to support measurement initiatives? Most papers seemed to assume that managerial support is absolute and constant- it either exists or not. This proves to be a false assumption in many cases. For instance, managers may easily lose their enthusiasm about performance measurement if they sense a lack of interest by elected executives and legislators (Ho, 2005). Management in such case may not be able to see any payoff for investing time and effort to collect and report performance information. Moreover, the systematic review indicates that internal politics is influenced by the perceived benefits of the system. It shows that both the adoption and the use of measures are affected by managerial choices which in turn are affected by the value of the measurement system. According to Hawke (2012), managers either choose between meeting minimum external requirements to supply performance information, or they see value in using this information for their own purpose. As performance measures improve efficiency, this will possibly lead to more accountability and better utilization of resources which might have a negative impact on the perceived benefit of measurement from the perspective of internal interest groups. This raises concerns about the political consequences of performance measurement which can be a major barrier to their adoption (Ho, 2005). For example, governments' efforts to lower budgets and cut expenditures as a direct response to the findings of particular measurement data might reduce incentives for top management to support performance measurement efforts (Cavalluzzo and Ittner, 2004). Managers are more likely to implement and use measurement innovation if they believe that such innovation is advantageous and can support their decision-making activities. Hence, the lack of a significant connection between the use of measures and management's willingness to implement change in some organizations could be explained in terms of managers' understanding about the value of performance data for making decisions (Folz, Abdelrazek and Chung, 2009). The systematic review suggests that there is a knowledge deficiency about this point. Future research should be helpful to form better understanding about the dynamism of internal political support and their perception on measurement initiatives. Future research should also examine the influence of internal politics on technical and cultural aspect of utilization.

External politics

External politics in this review relates to all forces outside a public organization that have an influence on the development and use of performance measurement. This includes the impact of all external stakeholders in power, as well as those who wish to exert political influence in opposition (Hawke, 2012). The literature indicates that external pressure is one of the main drivers for adopting performance measures in public organizations. It indicates that external stakeholders' interest in performance information will create strong pressure for the development of measures. For instance, many of performance measurement initiatives in governments have been credited largely to regulations such as the Competition Principle Agreement (CPA) in Australia, which mirrors similar regulatory initiatives in other countries in Europe and the USA (Pollanen, 2005). These initiatives show that external political support works as a catalyst for change by creating a demand for performance information (Tillema, Mimba and Helden, 2010). However, performance measurement that is promoted and adopted in response to pressure from external political elites is not automatically used (Lee and Kim, 2012). For example, a survey by Taylor (2011b) on a group of managers in state governments in Australia concludes that external support did not have a considerable influence. Authors explain that such support could be insignificant in cases where management is convinced that they cannot depend on stakeholders' support for the application of performance measurement in their agencies. Similar findings were reached by Folz, Abdelrazek and Chung (2009) who maintain that there is no statistically significant connection between the use of various types of performance information and the extent of support by elected officials. From his part, Sole (2009) argues that some sources of political influence, such as legal requirements to report performance, are more relevant to adoption than they are to the actual use of measures. Moynihan provides an interesting distinction between what he calls instrumental and symbolic value of the new system. According to him, for some external agents the symbolic benefits of the system are more important than the instrumental benefits. Performance measures therefore are adopted for institutional image and accountability purposes which indicate that the benefits would be symbolic and will not necessarily lead to improvement as suggested by external agents. External agents such as elected officials, according to him, are more likely to use the instrumental value of a reform program only as a selling strategy to accrue certain benefits such as positive image in the media or improving their re-election chances. This is what Van Dooren (2005) calls window-dressing cases in which a public organization may undertake many measurement initiatives but fail to actually use them. This implies that external politics have a partial influence with regard to adoption by creating the desire for change and not necessarily affecting change itself.

That is not to say that external pressure will necessarily lead to superficial use of PM in public organization in all cases. Instead, as implied by some papers, seeing the value of the measurement system is important to increase the likeliness of its application (Folz, Abdelrazek and Chung, 2009). For instance, some authors have stated that a key element of the actual use of performance measurement is the support from major external stakeholders such as citizens (Sole, 2009; Folz, Abdelrazek and Chung, 2009).

Citizen's interest in the result of measurement may place stronger pressure on decision makers to use performance information as a tool for communication and improvement (Ho 2005). Similarly, gaining the support of elected officials- *per se*- shouldn't be neglected given their power in legitimizing reforms, securing fund, and foreclosing back channels (Berman and Wang 2000). According to Yang and Hsieh (2007), elected officials may support adoption as a reform strategy on the one hand, but their desire to cut costs makes them hesitant to allow extra resources for actual initiatives on the other. Hence, external political pressure could in some cases have an influence on the actual use of measures even if it is in the negative direction.

The literature shows that one way to gain support from external interest groups is through involving them in the process of measurement. According to Ho (2005), the involvement of citizens and elected executives can certainly make a difference, as it gives a "political weight" to the tool and facilitate the integration of performance measurement in decision making. The general principle behind the idea is that more involvement will lead to more support. Conversely, less participation from elected officials in the design of performance measures may result in less support for its development and use (Berman and Wang, 2000). However, the literature indicates that this is not a conclusive finding. For example, Yang and Hsieh (2007) conclude by stating that although external stakeholders' participation is likely to enhance their satisfaction, it may not benefit internal management efficiency. Authors point to the organizational and individual obstacles to involving certain stakeholders such as citizens and elected officials in performance measurement. According to them, participation has a cost, implies the need for reforming traditional systems and processes, and requires supportive managers who are willing to share performance information and work with stakeholders. In order for involvement to have a positive impact over the utilization of performance measurement these conditions must be sustained first.

Examining the influence of external interest groups revealed several challenges that have significant implications. As discussed earlier, identification of key external stakeholders in public organization is not easy. Papers in the review varied in terms of defining one particular external source of influence for all public organizations. Various groups of stakeholders have been mentioned that include, but are not limited to, elected officials, citizens, labor unions, foreign and local funding bodies, legislators, and more broadly central or local governments (Julnes and Holzer, 2001; Yang and Hsieh, 2007; Berman and Wang, 2000; Moynihan, Pandey and Wright, 2011; Hawke, 2012; Inamdar, Kaplan and Reynolds, 2002; Mimba, Helden and Tillema, 2007; Tillema, Mimba and Helden, 2010). This indicates that external influence is not the same for all organizations. It shows that, as far as performance measurement is concerned, external stakeholders' influence varies according to the level of power of each and based on the context. Managing and meeting multiple and often conflicting requirements of various stakeholders is an additional challenge for effective utilization of measurement in the public sector which was not covered thoroughly by literature. This adds to the

complexity of studying the influence of external politics and generalizing the findings. Further empirical researches might be needed to examine the influence of specific external interest groups on the utilization process. Future research should also examine how that influence varies based on organizations' type or according to culture.

3.3.2 Technical Factors

3.3.2.1 Overview

The systematic review shows that there are several technical factors that need to be considered for successful adoption and use of performance measurement in the public domain. Technical factors are related to the available capacity for an organization to develop and use PMS. It includes the adequacy of available resources as well as the quality and appropriateness of the measurement design. According to Julnes and Holzer (2001), an emphasis on issues such as goal orientation, availability of resources, and access to information is a rational view that look at the utilization of measures from a technical perspective. The aim is to identify those rational factors that have implications for adopting and using performance measures. Several technical factors were cited by the literature. These were grouped and analyzed based on topic and relevance. Based on the content analysis, factors were associated with three main categories: measurement issues, availability of resources, and technical design of the measurement system.

Measurement issues

Performance is usually measured in terms of the level of achievement of organizational goals and objectives. A fundamental premise of many measurement frameworks is that the definition of indicators should be aligned with the vision and strategic objectives of the organization (Pereira, 2012; Kaplan, 2001; Kravchuk and Schack, 1996; Sá and Kanji, 2003; Zeppou and Sotirakou, 2003). The literature confirms the finding of Van Dooren (2005) and shows that measurability of services is a key factor for success and that organizations that have routine-based services have a higher adoption and use of performance measurement. However, the literature indicates that there are several challenges that limit the ability to measure performance in the public sector which negatively impact the success of measurement initiatives. It shows that there are many problems in relation to defining clear, attainable, and measurable goals in public organizations and indicates that in many public organizations, goals are usually undefined, vague, limited, generally broad and unachievable (Adcroft and Willis, 2005). Each of these factors will have a negative influence over an organization's ability to develop and use performance measures successfully. For example, due to the perpetual shift in political priorities, ambiguity could lead to dysfunctional consequences such as gaming and misrepresentation of results (Goh, 2012). This means that results could be subject to different interpretations and manipulation by users which, in effect, undermines the purpose of measurement. An additional challenge for measuring performance is related to the meaning and definition of 'performance' according to internal and external stakeholders. For example, Pollanen (2005) explains that performance in public organizations remains a relative concept that might mean different things for different users. According to him, the term is socially constructed by the multiplicity of stakeholders in the public domain and it is not likely that all aspects

of performance would be the same for all stakeholders. An example from health organizations is provided by Inamdar, Kaplan and Reynolds (2002) who explain that bearing in mind different needs for different type of customers, executives had to make difficult trade-offs to operationalize the different value propositions with measures that were aligned and did not conflict with each other.

Deciding what aspect of performance should be measured is an additional challenge and some papers referred to the difficulty of measuring all aspects of performance in the public domain. According to Carmona and Grönlund (2003), dimensions of police work that are important to the quality of urban life are missing due to the difficulty of measurement. Authors explain that concentration was placed on the easy- to-measure metrics such as number of arrests, clearance rates, and response time. In their opinion, these are measures that provided a traditional view of police work while some crucial areas were neglected. While the first scenario will pose additional challenge in terms of measurement, management and balance between multiple measures, the latter might provide incomplete or even distorted picture of performance *“it outlines an apparent contradiction between what police should do and what they actually do”* (Carmona and Grönlund, 2003, p.1483). A possible consequence is either proliferation of different kinds of measures or the easy to measure will drive away the more difficult (De Bruijn, 2007).

The review shows that there are some intangible assets that are difficult to capture and realize in the public domain (Northcott and Taulapapa, 2012; Cavalluzzo and Ittner, 2004). For instance, while it is relatively easier to measure the inputs and outputs of a particular services or programs, it is significantly more difficult to measure the outcomes (Sharma and Gadenne, 2011). The outcome of basic research and development activities are some of those tasks that are difficult to accurately measured using objective, quantifiable performance metrics (Cavalluzzo and Ittner, 2004). Further, measurement of some aspect of performance might take considerable time and efforts (Fryer, Antony and Ogden, 2009). Similarly, identifying and measuring long-term outcomes of many public services might not be attainable at all times (Pollanen, 2005). In such settings, a theoretical but possible feasible alternative would be an emphasis on subjective and qualitative judgments. The literature indicates that uncertainty in the identification of metrics and their use will continue to be one of the main concerns for managers in public organizations. It shows that there are many technical issues related to measurement systems in public organizations which need further investigation and research in order to ensure their success.

Resources availability

The literature confirms that a successful adoption and use of performance measurement is supported by the allocation of adequate resources. For some scholars, availability of resources is necessary for successful utilization (Lilian Chan 2004; Northcott and Taulapapa, 2012); others think that the availability of resources does not guarantee

success but inadequate resources guarantee failure (Moynihan, 2004; Rantanen et al., 2007). The systematic review shows that resources are needed during and after the development of measures. According to Moynihan (2004), considerable time and administrative attention should be devoted to the implementation of measurement systems. His cases show that resources' inadequacy could create a sense of crisis that would force organizations to focus on coping with immediate problems rather than on a long term strategic agenda.

Training is one component of resource which seems to affect both levels of utilization. Yang and Hsieh (2007) maintain that training improves rational cognition about performance measurement, hence is more relevant to the adoption level than the use level. Other literature informs us that if the design and development of measures require skilful and well-trained people it is equally important to train those people responsible for receiving and using information (Hawke, 2012). Training is not only necessary for facilitating adoption and use of measures but also a good indicator of the availability of an enabling culture that is based on orientation towards efficient goal achievement (Julnes and Holzer, 2001). Frequency of access to information related to the topic of performance measurement is a good measure of an organization interest in training (Sole and Schiuma, 2010).

The literature also shows that managerial capabilities are an essential component for the success or failure of any measurement initiative. According to Rantanen et al. (2007), the lack of managerial skills in Finnish public organizations is one of the main causes of problems. Authors explain that, in public organizations, public managers are not chosen based on their managerial skills but rather based on substance skills. This, according to them, means that an incompetent manager does not necessarily know what to manage let alone to measure.

With all that in mind, it should be noted that the literature was not explicit in terms of defining the type of resources that are needed to develop and use performance measurement. The term 'resource' is generally broad, unclear and could encompass all the needed requirements to achieve an end. Only few papers presented an overview of the meaning of resources and seem to point to general rather than specific elements such as the availability of finance, human resources, information systems, and time. The term should be used cautiously and future research should be explicit in terms of defining the aspects of the resources they refer to. Further, papers indicate that attainment of resources in public organizations is not easy most of the time. They indicate that the allocation of human and financial resources is a shared responsibility between internal and external stakeholders, mainly management and financing authorities. Aligning different and conflicting views of various stakeholders and getting their approval for a particular budget or programs is a considerable challenge that undermines the ability to utilize measurement efficiently. Also, as noted earlier, political actors allocate resources based on their perception of measurement value and according to their interests. If for instance, they realize that the application of new measurement system is going to undermine their authorities or negatively affect their interests they will be less likely

keen to support it. Thus, it seems reasonable to assume that availability of adequate resources could be a good indicator for the level of interest and support by political interest groups inside and outside an organization.

Technical Design

The systematic review indicates that both the supply and consumption of performance information are dependent on the quality of the technical design. Organizational performance success or failure can be determined by the application of appropriate performance measures (Cavalluzzo and Ittner, 2004). However, the literature notes that few attempts have been made to provide public organizations with PMS tailored explicitly to their needs (Micheli and Kennerley, 2005). Instead, various frameworks and practices have been traditionally imported to the public sector from the private one under the assumption that what works for the latter should inevitably work for the former. Papers denote that public sector organizations have distinct features which inevitably affect the design of PMS. As shall be discussed in the following section, the review has shown that for each of the frameworks that have been discussed there still many limitations that need to be considered before gauging their rightness for public organizations. Having said that, it is worth noting that the aim of this review is not to contribute to the discussion of various frameworks of measurement in the public domain, which is beyond the scope of this paper, but rather on the way that certain elements of the measurement design affects both adoption and use of measure. The review worked to form better understanding about the elements that make a specific design fit for purpose in public organizations. It corroborates the view that there is a need to understand the limitation of designs in terms of capturing the soft and hard measures to avoid situations where PM might misinform rather than inform (Micheli and Kennerley, 2005; Kravchuk and Schack, 1996). The analysis was limited to few papers that emerge from searching the databases and match the criteria as stated in the methodology section. Consequently, selected papers were mostly theoretical and centered on authors' experience and knowledge. The following section will begin by briefly exploring some of the frameworks thought to be working for public organizations followed by a discussion about the main characteristics that affect the utilization of PM in public organizations.

3.3.2.2 PMS Frameworks- General Overview

Balanced Scorecard

The Balanced Scorecard (BSC) introduced by Kaplan and Norton in 1992 aimed at providing top management with a comprehensive and balanced view about business. It was originally developed for private organizations to overcome some of the limitations of measurement that is centered on financial performance only. According to authors, no single measure can focus attention on the critical areas of the business and that management should not be forced to choose between two sets of measures financial and operational. Instead, the model aimed at integrating financial measures with some other operational measures on: customer satisfaction- how do customers see the organization; internal processes- what the organization need to excel at; and the organizations' innovation and improvement activities- can the organization continue to improve and create value. It was argued that by looking into business from these four perspectives, managers would align their activities to the main strategy and would be able to decide what is important and what needs to be done to improve performance. According to Inamdar, Kaplan and Reynolds (2002), the approach has the potential to deliver several benefits such as:

- Provide a framework for focus and alignment around strategy
- Establish core principles and process for strategy implementation
- Better communication and collaboration to assign responsibilities
- Better assessment for strategy progress
- Better utilization of resources towards the development of new products and services
- Continual feedback and adjustment to environment and needs

This approach gained exceptional recognition to the degree that it became difficult for any paper that discuss performance measurement to fail its mention, and hence it has become almost synonymous with PMS. Kaplan (2001) indicates that while the initial focus and application of the BSC was in the private sector, the opportunity to improve the management of public and non-profit organizations should be greater. He maintains that BSC has a great value for these organizations in a way that aligns all resources to accomplish organizational objectives. Given the multiplicity of objectives and stakeholders that characterize public organizations the BSC was considered by many as the appropriate tool that will direct managerial attention to more meaningful results by means of presenting multi-dimensional view of performance (Northcott and Taulapapa, 2012). This match between BSC and public sector management aims should have led to wider adoption and use in public organizations.

However, despite the growing number of articles that cover the BSC topic relatively very few were concerned with its application in the public domain (2012). Authors like Micheli and Kennerley (2005) indicate that many of those papers were written by advocates of the approach and there is still a need to conduct in-depth analysis of the issues that need to be overcome when it comes to implementation and use in the public sector. Furthermore, Sá and Kanji (2003) question whether the BSC can be regarded as a measurement model since it fails to identify specific variables and how they can be measured. He maintains that BSC is just a conceptual model that is not necessarily applicable to the needs and requirements of multiple stakeholders in public organizations. Kaplan and Norton (2001) recognize that applying their framework in the non-financial sector has some difficulties, on top of which are issues related to the definition of both strategy and customers. The issue becomes more complicated when considering the multiplicity of stakeholders and the share of production or decisions amongst public agencies and actors. Moreover, the framework does not explain how to draw a balance between various and even conflicting needs of different stakeholders. As such, modifying the architecture of the BSC was considered essential for its use in the public sector (Kaplan and Norton, 2001).

Performance Prism

The Performance Prism is another example of frameworks that were designed to fit the purpose of the private sector and were then thought to be useful for the public sector. Unlike the BSC, the approach developed by Neely, Adams and Crowe (2001) assumes that the starting point is the identification of both the organization's stakeholders and their needs. According to authors, the only reason for an organization to have a strategy is to deliver value to some set of stakeholders. They claim that their framework might be a second generation framework that addresses the shortcomings of many traditional measurement frameworks, including the balance scorecard. The framework stresses the need to develop scorecards that are appropriate to the business and consist of five interrelated facets: stakeholders, strategies, processes, capabilities and stakeholders' contribution. Consequently, the authors claimed that a wide variety of organizations, profit and not-for-profit, would be able to relate to and use the framework to improve performance. However, despite its potential, the model is yet to be tested and used in public organizations to determine its value and functionality.

The STAIR Model

A number of other performance measurement systems were originally developed for the public sector and, arguably, recognize the differences between public and private organizations. Sotirakou and Zeppou (2003) worked to examine the value of the STAIR model for measuring public sector performance by presenting the preliminary results of a pilot case study to modernize the Greek civil service. The authors suggest that organizational performance can be enhanced by the steps of the STAIR (strategy, targets, assessment, implementation, results) model, which stress "competence in

strategic thinking, strategic acting and strategic measuring, at all levels.” (Sotirakou and Zeppou, 2003, p.324) They argue that their model is not limited to measurement but is rather a comprehensive management framework that focuses on organizational success and builds lasting change. Similar to the BSC, the STAIR model emphasizes the development and clarification of an organizational strategy that is used for subsequent steps such as target setting and assessment of progress. According to authors, performance is a reflection of strategy and the main benefit from the STAIR model is that it bridges the gap between performance and strategic management process. The model proposition is: competence in strategic thinking, strategic acting and strategic measuring, at all levels, can enhance organizational performance. However, the model is yet to be tested in other public organizations in different contexts to gauge its applicability. It was built on the concept of strategic fit assuming that public organizations have clear, fixed and easily translatable strategies, which might not be the case at all times.

3.3.2.3 Characteristics of a successful design

For the purpose of the systematic review sub-question, characteristics of a successful design, and as inferred from the aforementioned frameworks and the content analysis of additional theoretical papers, there are specific features that characterize a successful PMS design. The following is a discussion about some of these features.

Alignment with Strategy

The main benefit of any measurement system is to translate organizational mission and strategy into a set of integrated measures (Sanger, 2008; Kaplan, 2001; Zeppou and Sotirakou, 2003). As can be seen from the aforementioned designs, the need for these designs to be aligned with organizational strategy is at the heart of the majority of designs. The goal as explained by Zeppou and Sotirakou (2003) is not only to improve performance in government organizations but, similar to the foundation of the BSC, to convert them into strategy-focused organizations. Kaplan and Norton (2001) and Kaplan (2001) explain that one important goal for using the BSC as a strategic management system is that it addresses serious deficiency of traditional systems which are not able to link long term strategic objectives with its short term actions. All levels of the organization would need to form a clear understanding about long term strategy and link their daily activities to it. Similarly, Kravchuk and Schack (1996) explain that performance measurement must start by formulating a clear understanding of the strategy and subsequent objectives which are the focal point for the development of any measurement approach. To them, an explicit measurement strategy is dependent on the formation of a clear and coherent understanding of an organization’s strategy, which then provides the blueprint for the design and development of the measurement system, including measures for input, output, and outcome. Failing to translate strategy into

specific targets, and accordingly into specific measures, undermines the validity of the measurement system. Additionally, an organization's interest in performance measurement could have an influence on its own strategies especially in cases where organization strategy is not completely clear or in need for reformation. According to Julnes and Holzer (2001), the process of measurement necessitate that goals need to be identified and targets need to be set. This process on itself will put organization's strategy into test. As such, the exercise of developing performance measures will force management to clarify the meaning of the strategy statement. However, the literature shows that public organizations face a real challenge in terms of defining their strategies and linking them to the overall performance. Uncertainties in the identification of objectives and the ambiguity in the meaning of 'performance' in the public sector along with problems associated with the multiplicity of stakeholders are some impediments that undermine the potential of any measurement system to align strategies with day to day work. Further empirical work should be encouraged to determine which framework is better able to achieve strategic alignment and under what conditions.

Stakeholders' Focus

Papers reviewed indicate that the success of any measurement design is determined by its value to stakeholders. According to Neely, Adams and Crowe (2001) the only reason an organization has a strategy is to deliver value to stakeholders. Such an approach is significant for governance and has wide implications on enhancing transparency, accountability and control (Sanger, 2008). The lack of a stakeholder-focused approach could also be responsible for information overload in public organizations (Jackson, 2005). According to Jackson, one could simply ask: "why some managers would have to receive a vast amount of operational information that they do not require or use but might be significant for some others?" The author maintains that this might explain why many public organizations have become keen to apply such approach.

However, the term 'stakeholders' is very broad, and measuring the value to or from various stakeholders might become problematic, particularly in the public sector. The literature shows that there is a challenge in identifying all stakeholders and it is not clear how a design could manage to draw a balance between various and even conflicting needs of different stakeholders. According to Kaplan and Norton (2001), in order for the BSC to work in public organizations there is a need for modification such as expanding the definition of customers in public organizations. They explain that in the private sector, customers both pay for and receive the service which is not necessarily the case in the public sector. Additionally, the literature indicates that accommodating the needs and requirements of a large number of stakeholders will be difficult and time consuming. There are many questions about the balance between competing or even conflicting goals. There is also the hazard of increasing the number of measures as the number of the organization's stakeholder's increases. According to Kravchuk and Schack (1996), the search of a single array of measures for all needs must be abandoned. Authors admit that this is a challenging and time consuming task and

recognize that performance may be subject to many conflicting interpretations due to the multiplicity of stakeholders in public organizations. However, authors argue that these challenges should be considered inevitable and normal part of the process which is necessary to lend credibility to the effort and avoid future misunderstanding. Further research is needed to learn more about the application of various designs in terms of their ability to incorporate stakeholders' requirement.

Balance

According to many papers in the review, 'balance' refers to the notion that performance is a multi-dimensional concept that implies the need to look into the organization as a whole by considering performance from different aspects instead of only one or few. It is better seen as an acknowledgement of performance complexity and to understand how different dimensions of performance inter-relate rather than limiting the view to one over-riding answer (Jackson, 2005). For instance, many authors were keen to highlight that traditional PMS doesn't seem to encourage continuous improvement and organizational learning, and should therefore be replaced by better designs that look into businesses holistically and by integrating financial and non-financial measures.

Sá and Kanji (2003) and Kaplan and Norton (2001) explain that traditional frameworks have important limitations that include being backward-looking, short-termism, control bias and lack of integration. They maintain that while private companies have recognized that financial measurements by themselves are inadequate for managing performance, the same principle is more applicable in public organizations. Jackson (2005), on the other hand, brings the attention to other dimensions of balance. According to him there are different overlapping elements of balance that need to be considered, particularly in public organizations. For instance, managers could be inclined to measure inputs and outputs, instead of outcomes and impacts, because they are easier to measure. Managers might also need to choose between measures of current versus future performance, long-term versus short-term targets, and balance between indicators of intended and unintended consequences. Managers need to balance multiple and to some extent conflicting requirements of various stakeholders. Furthermore, given the large number of stakeholders in public organizations, it is not clear how public organizations would be able to achieve the necessary balance between competing and sometimes conflicting needs. It is evident from the papers reviewed that balance is an issue for any design that will continue to be a challenge for public managers, and that further research is needed to better understand how they can achieve such balance.

Validity and Usefulness

For obvious reasons, the validity and usefulness of measurement comes at the heart for any design. Previous authors have indicated the advantages and value of their models using some empirical findings to support their claim. According to Sá and Kanji (2003), the validity and reliability of a PMS determines its success or failure. Papers reviewed

stress the value for money approach to determining the usefulness of measurement. That might be related to the cost in terms of money, time and effort that is required to develop and use PM in any organization before starting to reap the benefits. For Kravchuk and Schack (1996) a system of performance measures does no good if it does not inform decision makers or if it misrepresents, misleads or distorts information. A good example of such a case is given in the previous section about the move away from traditional performance measurement systems that were based on financial measures towards balanced and more valid systems that incorporate other operational measures such as customer satisfaction, internal process and organizational learning and growth.

However, the multiplicity of stakeholders and the co-production that characterize the public sector (De Bruijn, 2007) add to the complexity of gauging both the value and validity of measures, since different users will have different needs and are expected to have a different understanding and ways to utilize measurement. It should be noted that the validity and value of measurement systems is relative since they are determined by needs which continue to change over time according to changes in the internal and external environment, or changes of people and their requirements. Papers in the review did not say much about this and about the ways in which validity and usefulness of measurement can be assessed or the people who are responsible for valuation.

Clarity, Accuracy and Simplicity

For any specific design to be effective, it needs to be very clear and easy to use (Jackson, 2005; Sá and Kanji, 2003). The clarity and simplicity of measurement is an essential factor not only to gage the usefulness of a specific design but equally importantly the success and value of the measurement process. This can be achieved by defining all performance indicators in an open way and customizing measures to specific needs (Jackson, 2005). Jackson goes on to argue that clear, accurate and easy to use measures will affect attitudes by reducing dysfunctional behavior, and on the collection and interpretation of information. Clarity and accuracy, according to him, are also associated with the validity and reliability of indicators. Yet, simplification has a price too. For Kravchuk and Schack (1996), there is this challenge that in the face of performance information overload, decision-makers can be shielded from uncertainty and complexity through simplification and selective feedback information. However, the authors warn that oversimplification could have perverse effects as a consequence of little information reaching managers. Symmetry between simplification and dynamism in the design of PMS, according to authors, is therefore needed.

Participation and Ownership

A key element that is helpful for the success of PMS is the participation of stakeholders in the development of measures. This ownership suggests that PMS would need to be locally devised and applied (Jackson, 2005). Participation helps to avoid resistance and has its relevance to implementation and design (Taylor, 2011b; Sanger 2008; Rantanen et al. 2007). A traditional approach that is built on top down practice in terms of adopting measurement may be incompatible for numerous reasons (Sanger, 2008).

Sanger notes the difficulty for employees to attach meaning to measures that were developed by others and to accept the notion that their performance will be determined by measures outside their control. For Kaplan and Norton (1996), one essential approach to transforming BSC from a measurement tool into a more strategic one is through the process of communication and linking. As scorecards cascade down to individual level, measures are translated into objectives and measures that fit the requirements of all. The outcome is a clearer, more effective strategy available for everyone, and employees becoming more able to link their duties to the overall strategy. Additionally, involving others in the design and development of measures will be helpful to acquaint key decision-makers and others with the power as well as the limitations of the measurement system. While authors acknowledge that such process takes a lot of time and effort, they suggest it is an essential principle for more effective PMS.

It goes without saying that the issue of involvement and participation is directly related to the definition of relevant stakeholders who could participate in the design and development phase, which is problematic in the public domain. Participants might require a certain level of knowledge to participate effectively, and while previous models assume large participation is healthy, this remains theoretical since many papers in the review suggest that only a small group of people have actually participated in the development of measures. While participation is welcomed by reviewed papers, the issue could be more relevant to the context which makes generalization difficult and calls for further empirical research to reach a proper conclusion.

Dynamism

Measures lose their variability over time and shifts in strategic priorities or changes in the internal and external environment may dictate some modifications (Kravchuk and Schack, 1996). There is a need for some degree of flexibility in the sense of periodical revision of the measurement system in order to enable rapid adaptation to change. Jackson (2005) notes that performance measures require some time to develop and may require revision from time to time. The author warns against ignoring this issue by explaining that gaming is possible as people learn how to manipulate measures and use them for their personal interest as time passes. He associates dynamism with the value of measurement which requires new performance indicators to replace obsolete ones.

However, papers reviewed did not say much about the ways in which a performance measurement can be declared obsolete or may need change. Nor did they mention the responsibility for such decision. The issue of dynamism can be even greater when measurement is linked to the changing needs of stakeholders, as organizations would need to go into another round of accommodating the needs of these stakeholders. Further research is needed to learn more about how organizations determine the need for change in their measurement system and the value and cost of change.

3.3.3 Cultural Factors

The systematic review shows that measuring and managing performance in public organizations is more likely to contradict organizational culture and practice. It reveals that traditional cultural norms that are part of public bureaucracy may form fundamental and challenging obstacles (Sanger, 2008; Bevanda, Sinković and Currie, 2011; Bianchi and Rivenbark, 2012). Moving from classical models to a more strategy-based PMS will require the development of a new infrastructure, resolution of political problems, a shift in theoretical thinking, and re-education of the political elite (Šević, 2005). Yet, based on papers reviewed, relatively fewer studies have focused on cultural influence on the utilization of measurement in the public domain, compared with those concerned with political and technical factors. This supports the findings of Fryer, Antony and Ogden (2009) who argue that despite its significance, there are very limited researches concerned with the impact of culture on the development and use of measures.

Cultural elements take into consideration values inherent in the system and ways of operating, including societal, operational and institutional cultures (Hawke, 2012). According to Ho (2005), in order to make a difference, performance measurement needs to accompany changes in organizational culture and practices possibly by integrating performance information into daily operations and planning of the organization. The significance of culture over the utilization of measurement might be related to the time and efforts that are needed to induce change and to instill new values and practices in the public domain.

The development of measures requires considerable time, and even longer to change key actors' behavior from rejection to acceptance of the new approach (Torres, Pina and Yetano, 2011). As is the case with many change initiatives, adoption of measurement could be faced with a great level of resistance and skepticism. According to (Taylor, 2011a), an organizational culture that is resistant to performance measurement is one of the main challenges that affects the utilization of performance measurement in public sector. Taylor notes the perception of various respondents who see new system for measurement as an imposition thrust on them rather than a tool to help. Alternatively, he notes that a rational organizational culture is significantly associated with the use of performance indicators for decision-making. According to him, a rational culture would focus on specific elements that have some relation to performance measurement such as planning, goal-setting, and focus on end results. Several other authors were interested in the impact of an enabling culture on the utilization of measurement. A culture in which it is seen as a way of improving and identifying good performance and not a burden that is used to punish poor performers (Fryer, Antony and Ogden, 2009).

In order to identify the key cultural elements that influence the adoption and use of measurement in the public sector, the literature was content analyzed. Initial analysis shows that the basis of cultural influence consists of two main determinants:

performance-oriented culture and openness to change and learning. A performance-oriented culture could be defined as a culture in which the organization is clearly defining its goals and strategies, focusing on the end result, empowering employees, and seeking to collect information pertaining to its performance in order to improve it. Openness to change and learning relates to the perception and attitudes of management and employees towards new initiatives and their willingness to learn from past experience and about better ways to improve performance. It reflects an organizational motivation to build strong learning capacity and its commitment to encouraging evaluation and knowledge sharing (Hawke, 2012; Julnes and Holzer, 2001).

Other cultural considerations could be significant, especially in different contexts. For example, public organizations in developing countries are affected by certain, unique cultural challenges that are more relevant to their context than they are to other organizations in another countries. However, as noted earlier in the methodology section, the literature indicates that there are very limited numbers of studies on performance measurement utilization in non-western countries. The majority of papers in this review pertained to developed rather than developing countries. This makes generalization more difficult and raises concerns about the suitability of PMS for public organizations in non-western countries.

Performance-Oriented Culture

The literature informs that adoption and use of measures can be more successful in an agency that has certain capacities embedded in its organizational culture. According to Julnes and Holzer (2001), an emphasis on the need to develop performance improvement culture is key to successful utilization of measurement in public organizations. Authors like to promote the need for conducting an assessment of organization's readiness to develop and use measures prior to any decision or action concerning utilization. The reasons could be explained by Sole (2009), who maintains that public sector managers and employees very often feel threatened by performance measurement as it might force them to work harder and increase pressure to produce better results. According to him, certain elements sustain the role of performance-oriented culture as a factor affecting PMS in the public sector. Yet, while mentioned by some papers, a 'performance-oriented culture' and its components were not defined. My synthesis of relevant papers conclude that there are several functions characterize an organization with a performance-oriented culture: goal clarity, a focus on end results, collection and dissemination of performance data, and taking actions based on performance information (Sole, 2009; Mansor et al., 2012; Lee and Kim, 2012; Julnes and Holzer, 2001). The following are some of the main elements that describe a culture of performance measurement orientation as inferred from the literature.

Goal clarity: organizational culture determines the strategy and goals of an agency (Taylor, 2011b). The clarity of goals and strategies as well as understanding the link between them is one of the main considerations that have a positive impact on the two levels of utilization- adoption and use. Papers reviewed indicate that organizations that

are accustomed to defining their objectives and strategies clearly and communicate them to different organizational levels are better able to utilize measurement successfully. Clarity of goals makes measurement easier, together with the dissemination of performance information which eventually affects the use of measures. However, as noted earlier in the technical section, the literature informs that objectives and types of functions in public organizations are generally vague or even undefined. Unlike private companies, public organizations do not have a clear bottom line, such as profit, which can be anticipated and measured easily. Papers indicate that when goals are vague or unclear it is extremely difficult to establish how they can be accomplished, let alone to be measured. Undefined goals implicate different ways in which performance can be reported (Goh, 2012). They can create a behavior that is focused not necessarily on performance improvement but rather on showing compliance with numbers. Ambiguity can also lead to the dysfunction of measurement use such as gaming and misrepresentation. It cannot serve as a benchmark for performance assessment as it adds to the complexity of measuring progress objectively (Lee and Cho, 2011).

Focus on end result: this is a measure of the extent to which an organization is oriented towards efficient goal achievement. According to Julnes and Holzer (2001), this could be measured by: the extent to which programs and departments are guided by goals and objectives; the extent to which programs and departments communicate goals and objectives clearly; and whether there is agreement that an organization's mission promotes efficiency. However, in the experience of many public organizations, goals are not necessarily related to the overall strategy but rather a description of activities (Bianchi and Rivenbark, 2012; Rantanen et al., 2007). This lack of strategic focus or direction was thought to undermine the potential benefits of any measurement system. The study of Inamdar, Kaplan and Reynolds (2002) provides an example from several health organizations in the US that have attempted to utilize BSC. Authors note that, for many of those organizations that were unsuccessful to use the new system, the cause and effect logic of the learning and growth objectives in the BSC framework was difficult to learn, understand, employ, and consequently accepted mainly because people were not accustomed to thinking this deeply about the strategy. Similar finding was reached by Northcott and Taulapapa (2012) who maintain that since BSC design must flow from strategy a lack of a perceived strategic orientation may be a barrier to adoption. Hence, a communication of strategy to different levels of the organization deemed to be necessary for better organizational performance as well as for effective adoption and use of measures.

Rewards and incentives: performance-related rewards and incentives are a good demonstration of the use of performance information and commitment to the measurement system. Papers indicate that an obligation to produce detailed reports about achievements against pre-defined measures or indicators whether for internal or external stakeholders is a strong incentive to ensure that measures are actually developed and used. Similarly, rewarding and celebrating good performance, a good

representation of organizational culture (Sole and Schiuma, 2010), facilitates the introduction and deployment of measurement. Rewards and incentives could be used as a tool for building organizational capacity and inflicting change by motivating managers and workers to improve and measure performance (Sanger, 2008; Rantanen et al., 2007). From the review, the use of rewards and incentives aimed at changing behavior. It works as a catalyst for change in a way that makes change more favorable for stakeholders. While it is a necessary tool to mitigate resistance and gain support during the adoption phase, the incentive system is used as a tool to monitor and facilitate the use of information (Sole and Schiuma, 2010). However, whilst it is common practice in the private sector to motivate managers and employees alike by employing rewards and incentives as useful tools to improve performance the same is not applicable in public organizations. Linking rewards and incentives with performance is not possible at all times, especially when there is a difficulty in measuring performance in the first place (Sharma and Gadenne, 2011). The result would consequently vary according to the context, and based on other elements such as the availability of adequate resources and leadership support. Moreover, the lack of transparency and subjective reward criteria in public organizations is an additional concern that could have a negative impact, and limits the successful adoption and use of measures (Sotirakou and Zeppou, 2006).

Openness to change

The literature indicates that the application of a measurement system will cause a lot of changes and is more likely to affect people. It entails eruption of the traditional way for doing business which eventually will generate a great level of resistance that undermines the process of utilization. However, the literature indicates that an organizational culture that promotes innovation and is open to new ideas is more likely to succeed than others. It shows that organizations that continuously seek better ways to improve their performance have a better chance of accepting and using measurement successfully. According to Taylor (2011a), a major appeal of PM lies in the assumption that information is the basis of better decisions. Its application provides opportunities for decision-makers to learn about areas that are being measured, which will lead to better decisions.

A study on municipal governments in the USA and Canada concludes by suggesting that an organizational readiness for change is key to implementing BSC successfully. Authors like Lilian Chan (2004) and Julnes and Holzer (2001) suggest that organizational readiness for measurement can possibly explored by evaluating internal stakeholders' perceptions and attitudes toward change, and the presence of innovation reward systems. An evaluative culture that encourages knowledge-sharing and learning from feedback and performance results gives an indication about the possibility to adopt new initiatives (Goh, 2012; Taylor, 2011a). According to this view, organizations that are receptive to new ideas and willing to change their practices do not see change as a

threat but rather as an opportunity. They recognize the potential and value of change and consciously make a decision to embrace it. This will have a great influence on behavior in a way that is supportive towards the application of measurement.

However, while the literature seems to hypothesize a link between this factor and utilization of measurement in public organization, there seem to be very limited empirical research that examines the influence of this factor in reality. Furthermore, given the difficulty to establish a special system that rewards innovation, due to many limitations such as the lack of resources and the share of responsibilities by many stakeholders in public organizations, one would question whether it is really possible to assess an organizational openness to change. Taking into consideration that those who evaluate are the same as those who get evaluated for their views about change, questions of objectivity and validity are present. Management and employees alike might view themselves as supportive to change and seeking for better ways to improve, even if this is not the case in practice.

Other Cultural Considerations

The systematic review shows that there are more issues that need to be considered, particularly in cultures other than that in western countries. It indicates that public organizations in developing countries are affected by certain cultural challenges that are unique and more relevant to their context than they are to other organizations in another countries. For instance, the limited role of performance measurement in non-western countries according to (Mimba, Helden and Tillema, 2007; Tillema, Mimba and Helden, 2010) could be attributed to certain features that portray the public sector in these countries namely:

- low institutional capacity - inability of an institution to pursue its goals and improve performance;
- limited involvement of stakeholders- due to the sense of formality and lack of access to sufficient information;
- high level of corruption- misuse of power for personal gain that hampers public interest; and
- high level of informality – not following formal rules and regulations.

According to them, some of these factors can be found in all countries, but they are predominantly present in Least Developed Countries (LDCs). For instance, because of the low level income in LDCs, it is difficult to raise sufficient resources for an institutional capacity that promotes good performance. The weak control system in public organizations in these countries is an important source for corruption and informality. These features together with the local circumstances of each country are hindrances to the supply of performance information. That is, even if a demand for performance information does exist as a result of reform programs or intervention from stakeholders the supply is not guaranteed.

Based on their experience and reviewing a number of studies about public administration in LDCs, Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel (2006) present an overview of the unique challenges to effective reform in these countries. According to them, the introduction of an advanced framework such as the BSC might be difficult in LDCs due to many limitations that might dominate the culture of public organizations in these countries. Authors argue that despite the BSC potential to facilitate improvement, the lack of resources, politicization of public administration, and corruption are the major constraints undermining the BSC promise but dominant in these cultures. A similar conclusion was reached by Bevanda, Sinković and Currie (2011) who maintain that the assumption underlying PMS may be antithetical to the managerial culture of the developing countries. They conclude by stating that implementing PMS such as BSC in developing countries may encounter problems that are not encountered in a developed country.

Šević (2005) discuss the process of building performance measurement system in public organizations in a transitional country such as Serbia. According to him, many of the issues that have been resolved in Western countries have emerged as being problematic in the policy implementation in a constantly changing transitional environment. Šević makes an observation about the negative influence over former performance feedback instruments due to change in leadership where successors tried to discontinue initiatives launched by their predecessors. Author explains that governments have to decide to reinvent itself before reaping the benefits of PMS. Adopting PMS consequently requires the resolution of political problems, the development of a new infrastructure, and educating the political elite about the necessity to implement such systems. He also explains that performance measurement can only work in an environment dominated by New Public Management ideology that is characterized by de-politicization of the civil service and social professionalization as well as high technical competence of those responsible for the design and implementation of the new system.

However, it should be noted that others like Verheijen and Dobrolyubova (2007) claim that there are no major differences between countries in terms of their ability to utilize performance-based management systems successfully in the public sector. To them, other political and technical factors might be more relevant. Based on their analysis of public management reform process in two Baltic States (Latvia and Lithuania) authors conclude that an incremental, step-by-step approach to performance management systems' implementation, and dedicated reform team along with sufficient political support 'can help overcome systemic weakness and ensure successful implementation of technically complex reform' (Verheijen and Dobrolyubova, 2007, p.205).

The systematic review shows that the widely held notion that PMS are not suitable for developing countries cannot be authenticated due to the limited theoretical and empirical literature in this area. Much work has to be done to determine the influence of contextual factors on the adoption and use of performance information in developing countries.

3.4 Summary and Conclusions

This paper works to examine the conditions under which performance measurement could be utilized efficiently in the public domain. Determinants of successful utilization of performance measurement in public organizations in this paper were viewed in terms of two stages – adoption and use of measures. Adoption is about demand and supply of performance information. It starts by general recognition about value and benefits of a measurement initiative, and ends by the selection and development of measures. The second stage is related to the actual use of measures – i.e. consumption of performance information - for making decisions in public organizations.

Papers in the review indicate that the demand for performance information in public organizations is based on two main driving factors: an internal desire for improvement; and compliance with external pressure. As per Figure 16, the content analysis shows that both internal and external actors are motivated by the value they associate with measurement initiatives. When internal and external stakeholders realize the potential of the measurement system they will be more inclined towards its application. Conversely, they will be less likely to adopt the new system if it is going to undermine their power or negatively affecting their interest.

The reviewed literature demonstrates that performance information supply and consumption is affected by political, technical and cultural considerations. It indicates that the successful utilization of performance measures in public organizations is dependent on the role of political groups (stakeholders). Papers conclude by indicating that managerial support and leadership is a prerequisite before and after developing measures. According to these papers, if public managers realize the potential benefits of the measurement system, they will be more inclined towards its application and use. Equally important, the lack of managerial time and attention will contribute to the unsuccessful supply and consumption of performance information.

Receptive and supportive workers are thought to facilitate utilization. One way to achieve that is by involving workers in the development of measures. Additionally, while the literature indicates that external stakeholders' interest in performance information will create a strong pressure for the development of measures it provides little account about the influence of a particular external group(s) on the supply and consumption of performance information. This implies that external politics have a partial influence with regard to utilization in a way that it creates the desire for change and not necessarily affecting change itself.

Empirical and academic literature acknowledges the importance of technical factors in establishing and retaining PMS in public organizations. As noted by Sanger (2008), PMS are costly to develop, require considerable time and require significant investment in the design and training of staff for collecting, using, and reporting performance

information. The literature indicates that there are many limitations in relation to measuring performance that are innate to the public domain. Emerging technical issues that are related to the definition of goals, meaning of performance, and dimensions that need to be considered remain a dominant concern in the public sector.

This paper explained that, for the majority of public organizations, allocation of necessary human and financial resources is not automatic since the responsibility for making such decisions is shared between internal and external stakeholders. Both the supply and consumption of performance information are also dependent on the quality of the measurement design. The content analysis shows that a successful PMS design is characterized by its ability to: translate organizational strategy and objectives into specific measures or indicators, reflect the needs of stakeholders, and incorporate multidimensional perspectives. It should also be valid and fit for purpose, simple and easy to use, encourage stakeholder participation and dynamic that continue to evolve based on the needs or changes in the environment.

The literature also voices that adoption and use of measures can be more successful in agencies that are both open to change and have a performance-oriented culture. It shows that organizations that continuously seek better ways to improve their performance have a better chance to develop and use measurement successfully. Similarly, utilization of measurement can be affected by an emphasis on the need to develop a performance culture in which performance measurement is viewed as a vehicle for improvement (Julnes and Holzer, 2001). Moreover, the paper reveals that public organizations in non-western countries might be affected by certain cultural challenges that are unique and more relevant to their context. Problems of informality, low-institutional capacity and corruption are some of the features that might be dominant in developing countries which thought to inhabit the development and use of measurement.

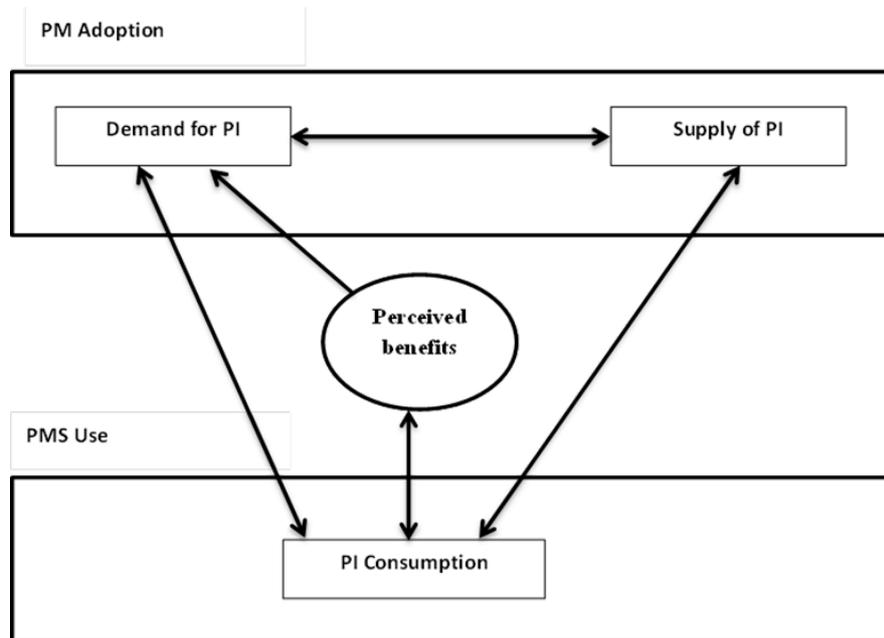


Figure 16 PI demand, supply and use

The systematic review unveils an interesting research agenda. Firstly, as I have shown in this paper, researchers have been looking at determinants of successful utilization of measurement, often without explicitly defining the meaning or the boundaries of the utilization process. For example, papers in the review have been less clear about the distinction between adopting and using performance measurement and, perhaps consequently, less conclusive regarding the determinants of both. This adds to the complexity of research in this field and inhibits its development.

Secondly, future research should be mindful of the dynamics of political support and their influence over the utilization of measurement. As noted in this paper both internal and external political groups are motivated by the value they associate with measurement initiatives. Their support therefore is not necessarily automatic and constant since the value from measurement could change over time as more performance information are supplied and used for decision-making. Moreover, the multiplicity of external stakeholders in the public sector continues to limit its ability to determine the influence of external politics in the utilization process. Further empirical researches that examine the influence of particular external interest groups on the utilization process should be encouraged. Given the low impact of external mandates on the actual use of performance information and the limited research on this area, further empirical studies are also needed to determine whether external political pressure will tend to be symbolic or with real influence on internal operations.

Thirdly, papers show that public sector organizations have distinct features which inevitably affect the design of PMS. The literature notes that few attempts have been made to provide public organizations with PMS tailored explicitly to their needs (Micheli and Kennerley, 2005). The review has shown that there are still many

limitations that need to be considered before gauging the suitability of a particular measurement system for public organizations. Hence, further empirical studies in this area should be encouraged to facilitate the integration and usefulness of PMS in the public sector.

Fourthly, fewer studies have been interested in the cultural influence compared to those concerned with political and technical factors. And even fewer seem to be interested on cultural elements in developing countries. Further research should be helpful to determining the influence of cultural factors on the success of measurement. It also could be helpful to determine whether measurement initiatives are suitable for cultures other than that in developed countries.

Finally, papers in the review seem to assume that political, technical, and cultural factors independently influence the development and use of PMS in public organizations. The majority of papers simply listed these factors but did not seem to study the interaction that might exist between them. However, as suggested earlier in this paper, these factors might interact to influence the development and use of measures. One or more factors or sub-factors can alter or accentuate the effect of others (Hawke, 2012). Hence, given the large number of potential interactions and limited theory on how these factors interrelate (Cavalluzzo and Ittner, 2004), future study should be mindful of the interrelationships that might exist between these factors and the significance of their interaction over the utilization of measurement.

In conclusion, the literature informs that utilization of performance measurement in the public sector continues to be a work in progress. Notwithstanding the growing interest on the conditions needed for adopting and using performance measures in the public sector, but the empirical work remains thin and fragmented, focused only on single issues, and lack an integrative perspective (Goh, 2012). For instance, there seem to be a knowledge deficiency in terms of the relationship between different levels of the utilization process. As per my synthesis Figure (16), the demand, supply, and consumption of performance information are related and influenced by the perceived benefits of measurement. More demand of performance information implies more supply and consumption. Similarly more supply might lead to more consumption. Further empirical researches are needed to narrow the gap in this field and to form better understanding about the interrelation between different levels of the utilization process in order to reap the benefits of measurement in the public sector.

Appendix

Authors and year	Title	Publication	Ranking	Type of Organization	Geography	Purpose
(Bevanda, Sinković, and Currie, 2011)	Implementing a performance measurement system in Croatia	Measuring Business Excellence	0	public utility-local government	Croatia	To describe the problems encountered when a performance management system such as the balanced scorecard (BSC) is implemented in the culture of a developing country (Croatia).
(Kaplan, 2001)	Strategic Performance Measurement and Management in Non-profit Organizations	NONPROFIT MANAGEMENT and LEADERSHIP	2	Social service organization	USA	Describes the adaptation of a new performance measurement and management approach, the

						Balanced Scorecard, to the non-profit sector. Several examples of actual implementation are provided.
(Liu, 2009)	Implementing and Evaluating Performance Measurement Initiative in Public Leisure Facilities: An Action Research Project	Systemic Practice and Action Research	1	three public leisure centres	UK	Addresses issues when implementing and evaluating performance measurement initiative in Leisure facility.
(Neely, Adams, and Crowe, 2001)	The performance prism in practice	Measuring Business Excellence	0	private organizations and a charity organization	UK	Describes and illustrates the practical application of a new measurement framework, the Performance Prism, which addresses the shortcomings of

						many of the traditional measurement frameworks being used by organizations today
(Pereira and Melão, 2012)	The implementation of the balanced scorecard in a school district: Lessons learned from an action research study	International Journal of Productivity and Performance Management	0	school district	Portuguese	To investigate the benefits, obstacles and challenges in implementing the balanced scorecard (BSC) in non-higher education public schools, more specifically, in a Portuguese school district.
(Rantanen et al, 2007)	Performance measurement systems in the Finnish public sector	International Journal of Public Sector Management	1	university, state agency, maintenance function	Finland	Aims to identify the specific problems faced by the Finnish public

						sector organizations in designing and implementing performance measurement systems (PMS)
(Carmona and Grönlund, 2003)	Measures vs. actions: the balanced scorecard in Swedish Law Enforcement	International Journal of Operations and Production Management	3	Law Enforcement	Sweden	Study examines the deployment of the balanced scorecard in police work in Sweden.
(Bianchi and Rivenbark, 2012)	A Comparative Analysis of Performance Management Systems The Cases of Sicily and North Carolina	Public Performance and Management Review	1	regional governments	USA	To contribute to the maturity process of performance management in regional governments, drawing conclusions primarily from the strengths and limitations

						of existing models in order to identify specific strategies on how to develop more robust performance management systems for higher levels of accountability and transparency in a time of global economic crisis
(Radnor and Lovell, 2003)	Success factors for implementation of the balanced scorecard in a NHS multi-agency setting	International Journal of Health Care Quality Assurance	0	Health Action Zone	UK	Aims at conditionally outlining some of the grounds that support the application of BSC in National Health Service (NHS). It then outline some of the factors

						for successful implementation
(Torres, Pina and Yetano, 2011)	Performance Measurement in Spanish Local Governments. a Cross-Case Comparison Study	Public Administration	3	Local governments	Spain	The nature of, and extent to which, performance measures are developed and how they are used in some of the biggest Spanish local governments. Why non-mandatory performance indicators are introduced and why specific initiatives are undertaken .
(Moynihan, 2004)	Why and How Do State Governments Adopt and	Journal of Public Administration Research and Theory	2	state governments	USA	Proposes a theory of adoption and implementation of

	Implement "Managing for Results" Reforms?					managing for results policies
(Sanger, 2008)	From Measurement to Management: Breaking through the Barriers to State and Local Performance	Public Administration Review	4	states and local governments	USA	Works to find some issues to understand how well citizens are engaged in either development or consumption of performance measurements. The paper also explores the determinants of successful implementation, adoption, and use of performance measurement.

(DeGroff et al., 2010)	Challenges and strategies in applying performance measurement to federal public health programs.	Evaluation and program planning	0	federal level	NA	To explore y several challenges impede developing and implementing performance measurement systems at the federal level and offer potential solutions that support the development of robust and practical performance measures to meet the needs for program improvement and accountability
(Fryer, Antony, and Ogden, 2009)	Performance management in the public sector	International Journal of Public Sector Management	1	NA	NA	To assess the state of performance management within the public sector and suggest

						areas for further research.
(Goh, 2012)	Making performance measurement systems more effective in public sector organizations	Measuring Business Excellence	0	NA	NA	Argue that the questionable benefits and many barriers, challenges and problems with implementing performance management and measurement in the public sector environment are all due to the lack of focus on the process of managing the implementation of performance measurement.
(Kravchuk and Schack, 1996)	Designing Effective Performance-e-	Public Administration Review	4	NA	USA	Introduce fundamental design principles

	Measurement Systems under the Government Performance and Results Act of 1993					for engineering effective performance measurement system.
(Micheli and Kennerly, 2005)	Performance measurement frameworks in public and non-profit sectors	Production Planning and Control	2	NA	UK	To review the frameworks currently developed and implemented in public and non-profit organizations and to identify the requirements of a framework, which can be applied in this context
(Sole, 2009)	A management model and factors driving performance in public organizations	Measuring Business Excellence	0	NA	NA	Analyze the characteristics of performance measurement and management systems in public

						organizational and to highlight the main factors driving performance in this sector.
(Tillema, Mimba, and Helden, 2010)	Understanding The Changing Role Of Public Sector Performance Measurement In Less Developed Countries	public administration and development	2	NA	NA	This article develops a framework for understanding changes in the demand for and supply of performance information in public sector organizations in less developed countries. New Institutional Sociology (NIS) is used to argue that pressures from specific stakeholders stimulate

						organizational information. The
(Folz, Abdelrazek, and Chung, 2009)	The Adoption, Use, and Impacts of Performance Measures in Medium-Size Cities: Progress Toward Performance Management	Public Performance and Management Review	1	cities	USA	Examines the patterns of adoption, use, and impacts of performance measures for the purpose of advancing understanding of the challenges involved in moving from performance measurement to performance management.

(Kwon and Jang, 2011)	Motivations Behind Using Performance Measurement: City-wide vs. Selective Users	Local Government Studies	3	local governments	USA	Identifies and describes the utilization of performance measurement in Florida local governments. Two research questions are considered here: which factors influence Florida city governments to use performance measurement? And how do the factors associated with cities that are city-wide users differ from cities that are selective users?
-----------------------	---	--------------------------	---	-------------------	-----	--

(Poister and Streib, 1999)	Performance Measurement in Municipal Government: Assessing the State of the Practice	Public Administration Review	4	Local municipalities	USA	Examine the extent to which performance measurement become integrated into contemporary local government management
(Sotirakou and Zeppou, 2006)	Utilizing performance measurement to modernize the Greek public sector	Management Decision	1	government departments	Greece	Presents the results of a qualitative (focus groups) and quantitative (questionnaires) research approaches undertaken in the Greek public administration with newly appointed administrative staff in various governme

						nt departmen ts. The STAIR model was used as a conceptual tool for critical reflection on the issue of performan ce manageme nt and measurem ent.
(Inamdar, Kaplan, and Reynolds , 2002)	Applying the balanced scorecard in healthcare provider organizatio ns	Journal of Healthcare Management	0	healthcar e organizat ions	USA	Work to evaluate the potential of BSC as a strategic manageme nt tool for healthcare organizati ons. It identifies important factors for adoption and implement ation, and present challenges and barriers of the BSC application . It also

						works to introduce a proper guideline for BSC implementation.
(Sá and Kanji, 2003)	Finding the path to organizational excellence in Portugese local government: A performance measurement approach	Total Quality Management and Business Excellence	1	municipality	Portugal	Presenting a new performance measurement system that, according to authors, is holistic, comprehensive, valid, reliable and easy to use, and reports the results of its application to the Portuguese municipalities.
(Northcott and Taulapapa, 2012)	Using the balanced scorecard to manage performance in public sector organizations Issues and challenges	International Journal of Public Sector Management	1	local government	New Zealand	To examine the use of the balanced scorecard (BSC) as a performance management tool in the public

						sector. It seeks to identify issues and challenges in implementing the BSC in public sector contexts.
(Mansor et al., 2012)	Organizational Factors Influencing Performance Management System in Higher Educational Institution of South East Asia	Procedia - Social and Behavioral Sciences	0	University	Malaysia	To identify the factors that are influencing the implementation of performance management system (PMS) in South East Asia. This study specifically inspects organizational factors influencing PMS.
(Lee and Cho, 2011)	In search of a better fit: Exploring conditions for successful	International Review of Public Administration	0	federal agencies	USA	Examines the conditions for successful implementation of

	performance-based human resources management in U.S. federal agencies					performance-based human resources management (PHRM) in U.S. federal agencies.
(Sharma and Gadenne, 2011)	Balanced Scorecard Implementation in a Local Government Authority: Issues and Challenges	Australian Journal of Public Administration	0	Local Government Authority	Australia	Investigates whether the implementation of a BSC has been of value to a large Local Government Authority (LGA) within Australia. The research takes a thematic approach to identify the issues, challenges and lessons learnt in relation to the design and implementation of

						the BSC
(Hawke, 2012)	Australian public sector performance management: success or stagnation?	International Journal of Productivity and Performance Management	0	NA	Australia	To identify and explain the key factors affecting the success of Australia's public sector performance management system.
(Cavalluzzo and Ittner, 2004)	Implementing performance measurement innovations: evidence from government	Accounting, Organizations and Society	4	Government agencies	US	Examine some of the factors influencing the development, use, and perceived benefits of results-oriented performance measures in government activities.

(Lilian Chan, 2004)	Performance measurement and adoption of balanced scorecards A survey of municipal governments in the USA and Canada	International Journal of Public Sector Management	1	Municipal governments	US and Canada	The paper explores the results of a survey of municipal administrators regarding their organization's experience with performance measurement system as well as the balanced scorecard.
(Van Dooren, 2005)	What Makes Organizations Measure? Hypotheses On The Causes And Conditions For Performance Measurement	Financial Accountability and Management	0	Ministry	Belgium	Identify and study factors that drive measurement in public organizations. Six hypotheses on why organizations measure are tested
(Moynihan, Pandey, and Wright,	Setting the Table : How Transformational	Journal of Public Administration Research and Theory	2	Local Governments	USA	Offers a theory of how leadership affects the

2011)	Leadership Fosters Performanc e Informatio n Use					implement ation of manageme nt reforms. This article develops a theoretical model that proposes that transforma tional leaders “set the table” for performan ce informatio n use via a positive but indirect effect on two mediating factors, goal clarity and organizati onal culture.
(Pollanen , 2005)	Performanc e measureme nt in municipalit ies Empirical evidence in Canadian context	International Journal of Public Sector Management	1	municipa lities	Canada	Examines the actual and desired use of performan ce measures for manageme nt and

						external reporting purposes, as well as perceived impediments to their effective use.
Berman and Wang, (2000)	Performance Measurement in US Counties : Capacity for Reform	Public Administration Review	4	counties	USA	Examines the capacity of U.S. counties to undertake performance measurement.
(Julnes and Holzer, 2001)	Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation	Public Administration Review	4	state and local government	USA	Examines the factors that affect the utilization of performance measurement.
(Sole and Schiuma, 2010)	Using performance measures in public organizations:	Measuring Business Excellence	0	municipalities and local governments	Italy	To investigate the main challenges related to the use of

	challenges of Italian public administrations					performance measures in public organizations in order to identify the factors affecting the adoption and implementation of performance measurement systems.
(Yang and Hsieh, 2007)	Managerial Effectiveness of Government Performance Measurement: Testing a Middle-Range Model	Public Administration Review	4	district governments	Taipei, Taiwan	Fills the void by providing a middle-range theory and testing how external political support influences aspects of performance measurement adoption and effectiveness. It uses structural equation

						modelling to test a model that assesses how political environment, stakeholder participation, organizational support, and training affect the adoption and managerial effectiveness of performance management
(Jackson, 2005)	Falling from a Great Height : Principles of Good Practice in Performance Measurement and the Perils of Top Down Determination of Performance	Local Government Studies	3	NA	NA	Uses an evaluator's perspective and practical experience of implementing performance measurement systems in the public, private and

	e Indicators					voluntary sectors to devise ten principles of good practice in performance measurement
(Zeppou and Sotirakou, 2003)	The “STAIR” model: A comprehensive approach for managing and measuring government performance in the post-modern era	International Journal of Public Sector Management	1	civil service	Greece	Introduces the preliminary results of the STAIR
(Behn, 2003)	Why Measure Performance? Different Purposes Require Different Measures	Public Administration Review	4	NA	NA	To present a theoretical foundation for measurement by discussing the purposes of measurement within the public sector

(Kaplan and Norton, 2001)	Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part I	Accounting Horizons	3	NA	USA	Explain the advantage and features of BSC use in private and public organizations
(Adcroft and Willis 2005)	The (un)intended outcome of public sector performance measurement	International Journal of Public Sector Management	1	NHS, Higher Education	UK	To consider the extent to which regimes of performance measurement in the public sector are fit for purpose, and the likely outcomes for public services and public sector workers of such performance measurement systems.

(De Bruijn, 2002)	Performance measurement in the public sector: strategies to cope with the risks of performance measurement	International Journal of Public Sector Management	1	NA	NA	Presenting five strategies to prevent the perverse effects of PMS
(Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006)	Reforming Government Agencies Internationally : Is There a Role for the Balanced Scorecard ?	International Journal of Public Administration	1	Local Government	USA, UK, Sweden, New Zealand, Australia	After discussing domestic and foreign use of the Balanced Scorecard in government agencies in developed countries, this article explores the applicability of the model to government organizations in developing countries.

(Mimba, Helden and Tillema, 2007)	Public sector performance measurement in developing countries: A literature review and research agenda	Journal of Accounting and Organizational Change	0		NA	To explore the influence of specific characteristics of the public sector in developing countries (i.e. A low-institutional capacity, a limited involvement of stakeholders, and high levels of corruption and informality), and of reforms of this sector, on public sector performance measurement.
(Moynihan and Pandey 2010)	The Big Question for Performance Management : Why Do Managers	Journal of Public Administration Research and Theory	2	local government	USA	Examines the antecedents of self-reported performance information use from

	Use Performance Information ?					a survey of local government managers.
(Perera, Schoch and Sabaratnam, 2007)	Adoption of the balanced scorecard in local government organizations : an exploratory study	Asia-Pacific Management Accounting	0	local Government	Australia	Adds to the existing literature on the BSC by examining its use in local governments. The study examines the extent of adoption of the BSC in local governments and the factors that appear to influence its adoption.
(Šević 2005)	Measuring performance on a local government level in a transitional country: the case of Serbia	International Journal of Public Sector Management	1	local governments	Serbia	Discuss the process of performance management and measurement system-building in the public

						sector in a transitiona l country, with particular reference to Serbia
(Verheije n and Dobrolyu bova 2007)	Performanc e managemen t in the Baltic States and Russia: success against the odds?	International Review Of Administrative Sciences	2	NA	Latvia, Lithuania, Russia	The introductio n of such reforms can be successful even in public managemen t systems that are not ‘advanced’ . This contradicts the widely held notion that performan ce-based public managemen t systems are not suitable for ‘developin g’ countries.
(Ho, 2005)	Accounting for the Value of Performanc e Measureme	Journal of Public Administration Research and Theory	2	NA	USA	Studies the decades-long question, “Does performan

	nt from the Perspective of Midwestern Mayors					ce measurement matter?’’ and discusses the implications for future results-oriented reforms
(Taylor, 2011a)	Strengthening The Link Between Performance Measurement And Decision Making	Public Administration	3	Various public orgs.	Australia	whether the demands that are made on public agencies to externally report on their performance have a limited Or wide ranging impact on the utilization of performance indicators (PIs) for internal improvements. It also

						examines the main factors that affect the use of PIs for decision-making.
(Taylor, 2011b)	Factors Influencing The Use Of Performance Information For Decision Making In Australian State Agencies	Public Administration	3	Various public orgs.	Australia	Is the state of a performance measurement system the most important element for promoting the Utilization of performance indicators (PIs) in the public sector? Or are there other more influential factors, such as organizational culture, or even individual perceptions on the merit of performan

						ce Measurement for their agency?
--	--	--	--	--	--	---

**CHAPTER 5: THE PATTERNS OF PERFORMANCE MEASUREMENT
APPLICATION IN THE YEMENI PUBLIC SECTOR**

Introduction

The field of performance measurement and management has attracted great level of attention in the literature over the past decades. Performance Measurement and management systems have become important for reforming private and public organizations in many places in the world. Proponents of Performance Measurement Systems (PMS) claim that an effective utilization of measurement will add value to the public service by enhancing quality, increasing transparency and accountability, and making public organizations more sensitive to the need of people. However, despite their potentials, systems of performance management and measurement did not seem to be a prime component of several reform initiatives in the public sector or to the very best have been contained implicitly within past initiatives with no explicit concern of their influence on performance improvement (Van Dooren, 2005). Researches on whether the utilization of PM would mean better performance and outcomes, particularly in developing countries, remain thin with evidence of both success and failure to integrate measurement into public management (Gao, 2015; Goh, 2012).

This paper presents an empirical research for Project 2 as part of the Cranfield DBA program. It is an exploratory study that seeks to establish practices of Performance Measurement (PM) adoption and use within public organizations in the Republic of Yemen. The research worked to obtain and examines the views of several public officials on the actual and potential role of PM to support reform and improve performance in the public sector. Given the limited research on public reform programs in developing countries compared to similar initiatives in developed countries (Tillema, Mimba and Helden, 2010; Gao, 2015), this study provides an initial insight into the existing situation in order to narrow the gap in knowledge about the utilization of Performance Measurement System (PMS) in the non-developed world. The paper reflects the significant importance of the context and subsequent processes and arrangements that promote or impede the development and use of PM as a reform tool in public organizations.

Research Questions

This paper cast the following question: what are the patterns of PM usage in the Yemeni public sector? It pursued answer by relying on the views and experience of senior public officials and managers in Yemen. The research was guided by three main sub-questions that were developed to fill the gap in the current literature on the use of PM in the public sector:

- How do public managers perceive the use and usefulness of measurement?
- What are the types and sources of demand and supply of Performance Information (PI) in public organizations?
- What are the main barriers to the adoption and use of measures?

Literature Review

Overview

The efficiency and effectiveness of public services have been under great scrutiny by many stakeholders including citizens and decision-makers. The traditional and classical approaches to public management, it was thought, need to be replaced with new techniques that improve the value of public products. According to Gruening (2001), the new discipline drew on the findings of the neoclassic researchers and witnessed the transplant of different approaches from the private sector. For instance, the progressive reforms in the United States during the 1920s were based on various principles such as unity of command, hierarchy, specialization and division of work, and delegation of authority (Adcroft and Willies, 2005; Fábíán, 2010). The advent of New Public Management (NPM) have led to the creation of a new rational style of public administration in an endeavor to present a framework that is applicable for all organizations at all times and is capable to solve all the problems of governments and bring change towards more efficient and effective management (Hood, 1991; Hood, 2000). This new rational style was marked by a heavy reliance on information-gathering as a means to providing the best solutions to management problems, and an emphasis on the objective measurement of performance. Zero-base budgeting, public sector marketing, management by objectives, strategic management and performance measurement are all examples of the rational thinking that dominated since the mid of the past century.

The principles of measurement are not new to the management field. According to Khan and Shah (2011), performance measurement has its roots in accounting since the middle ages, when it was used to settle transactions between traders. Holzer and Kloby (2005) is with the view that PM as we know it today can be traced as far back as 1912 when New York Bureau of Municipal Research used the budget for showing performance of government rather than accounting purposes only.

Following the introduction of NPM reforms in many OECD countries, the idea of measuring and managing performance has been on the agenda of many public organizations that seek to improve efficiency, minimize waste and provide value for money services (Micheli and Manzoni, 2010). The potential impact on organizational goals has encouraged many to use performance measures as an effective tool to achieve several ends. For instance, performance measurement can be used to: change behaviors, monitor policy goals, hold managers accountable for outcomes, provide basis for planning and control, improve performance and management practices, creates focus for the future, and enhance communication and understanding between different levels in an organization (Behn, 2003; Gianakis, 2002; Fryer, Antony and Ogden, 2009; Jackson, 1988).

Performance Measurement and Performance Management

Many scholars use the terms performance measurement and management interchangeably with little consideration for the difference between the two (Goh, 2012; Fryer, Antony and Ogden, 2009). According to Van Dooren (2005), performance measurement is a fuzzy word that has different meanings to different people. Fryer, Antony and Ogden, (2009) define performance measurement as quantifying the input, output, or outcome of a particular operation, services, process, activity, program or even policy. For Pollanen (2005), PM is an important device for planning, control, and decision-making that is concerned with the assessment of achievement of organizational and managerial objectives. It is meant to objectively determine the achievement of an organization to meet its objectives and contribution to the well-being of its wider community. Degroff et al. (2010) and Elg and Kollberg (2009) consider it the regular process of defining, monitoring and using of objective indicators, to determine the overall performance of an organization or program. Typical categories of measures, according to them, include process measures such as inputs and outputs, and outcome measures with immediate, intermediate, or long term focus.

Performance management, on the other hand, is viewed as a management tool to enhance the overall performance of an organization (Goh, 2012). It is about action rather than quantifying (Fryer, Antony and Ogden, 2009). It include all the “interrelated strategies and activities to improve the performance of individuals, teams and organizations” (Hawke, 2012, p.310). Performance management expected to specify goals and targets, allocate decision rights, and monitor and measure performance (Verbeeten, 2008). Optimizing performance and ensuring strategic evolution and goal congruence cannot be achieved without a system for performance management (Lilian Chan, 2004).

The distinction between performance measurement and management brings to the light similar distinction between adoption and implementation/actual use of measures. However, only few papers make a distinction between these two terms. According to Van Dooren (2005) and Julnes and Holzer (2001), adoption is about development of a capacity to act based on performance measures, i.e. whether an organization has or has not performance measures. Implementation, on the other hand, is about the impact of measurement or simply the extent to which the organization is making use of measurement information (Ammons and Rivenbark, 2008). It is “a question of how performance measurement is fed into the administrative and political system” (Johnsen 2005, p10). Sole (2009) presents a useful illustration to better understand performance measurement and management in terms of adoption and use of measures. According to him, while measurement represents the process of adopting measures, performance management refers to the actual use of these measures. The author explains that such distinction is useful to understand the factors that affect performance measurement and management.

Mimba, Helden and Tillema (2007) from their part wanted to avoid confusion about the exact meaning of terms like ‘development’, ‘adoption’, ‘implementation’, and ‘use’ of PM by introducing the supply/demand approach to study PM in the public sector. The supply of performance information refers to the production of measurement information which will be based on purposeful selection of performance areas, development of indicators, and collecting, analyzing and reporting performance information. The demand for performance information is viewed in terms of the actual or desired use of that information for making decisions either by managers and policy makers or by any other stakeholders who show interest in learning more about the performance of public organizations. There is growing demand that would require more supply of performance data which could lead to the adoption of PMs (Mimba, Helden and Tillema, 2007; Moynihan, Pandey and Wright, 2011; Tillema, Mimba and Helden, 2010; Torres, Pina and Yetano, 2011). Noting the difference between the two, Yang and Hsieh (2007) and Taylor (2011b) are with the view that demand for measurement information is a reflection of public organizations’ concern about accountability, responsiveness, and quality of service.

But why is it important to differentiate between performance measurement and management or between adoption and implementation of PMS?! A possible answer could be provided by Julnes and Holzer (2001) who explain the importance of making such distinction by indicating that recognizing these stages is helpful to identify and isolate the effects of factors for each stage which would support effective utilization of PMS. Authors call for conducting an assessment for organizational readiness to develop and implement PM. According to them, this will form a better understanding about the level of knowledge in the organization about the usefulness of measurement information, the level of support for measurement, and existing culture, resources, and expertise. Further, adoption and implementation may have different determinants as suggested by Yang and Hsieh (2007). Authors explain that while adoption may be driven by external pressure, actual implementation is affected by managers who could acquiesce, avoid, manipulate, or even defy the use of performance measurement in their organizations. Furthermore, the mere development of measures does not guarantee their use leading to the distinction between the availability of performance information and its usefulness (Taylor, 2011b; Berman and Wang, 2000; Torres, Pina and Yetano, 2011).

The existing literature suggests that performance measurement is not enough on its own and there is a shift in discussion to the broader and supplementary approach of performance management (Gao, 2015). Performance measurement may provide the information base that feeds into decision making about best ways to achieve desired results, but managing performance is about the actual use of measures to achieve desired goals and improve performance. According to Van Dooren (2005), a policy for performance management would need to follow a policy for measurement and not before.

Institutionalizing Performance Measurement

The introduction of PMS as a tool for managing performance in the public sector has been largely credit to different regulations in developed countries. In the USA, many government initiatives were taken to promote PM including the Chief Financial Officers Act, the National Performance Review, and the Government Performance and Results Act (Cavalluzzo and Ittner, 2004; Newcomer, 2007). These initiatives require first the clarification of missions and objectives of each government's project, program, or operation and second evaluating performance by measuring relevant outputs and outcomes. The British government worked to improve National Health Service using the Performance and Assessment Framework as a strategic and balance tool for monitoring progress. The Local Government Acts on the other hand was used to classify local governments into four categories ranging from high performers into poor performers based on assessments of performance criteria for local authorities (Chang, 2007; Micheli and Kennerly, 2005). In Australia, the advent of PM in local governments has been credit to various regulations by different states in response to the Competition Principle Agreement (Pollanen, 2005).

While proponents of PM assumed that legalizing and mandating reporting of results-oriented performance measures can improve efficiency and effectiveness by making government and public managers more accountable for results, many studies suggest that integrating PMS into the culture of public organizations is rather complex (Ohemeng, 2011; Cavalluzzo and Ittner, 2004; Van Dooren, 2005; Hawke, 2012; Torres, Pina and Yetano, 2011; Tillema, Mimba and Helden, 2010; Mimba, Helden and Tillema, 2007). Public organizations may satisfy external requirement for measurement by focusing only on developing measures of performance without necessarily using the information for decision making or performance improvement purposes. In such case, the symbolic benefits of measurement will be more important than the instrumental benefits and PM will be used, mainly, to legitimize performance to stakeholders in order to gain certain advantages.

PM in Developing Countries - A Research Agenda

A quick look into the literature reveals that the field of performance measurement and management has been growing over the past decade with much to learn from the experience of many organizations around the world. However, there is significantly smaller number of empirical researches in the public sector than that in the private sector, and most empirical and theoretical papers published in academic journals have been largely confined to western countries, mostly the USA and UK, with little attention to other contexts such as that in developing countries (Mimba, Helden and Tillema, 2007; Taylor, 2011b; Gao, 2015). The literature review denotes that adjusting PM to the specific cultures in developing countries continue to be a challenge and there is a need

for further research to understand the link between national and organizational culture and possible difficulties encountered while attempting to implement PMS (Bevanda, Sinković and Currie, 2011; Šević, 2005; Tillema, Mimba and Helden, 2010; Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006; Nath and Sharma, 2014).

To that end, authors like Van Dooren (2005) and Siddiquee (2010) suggest that development of PM policy should be based on learning more about the organizational characteristics that foster such policy, and warns that the value and benefits of NPM principles- including measurement- cannot be taken for granted particularly in developing countries. Authors explain that since reforms have been limited in developing countries, advocates of performance measurement and management principles should acknowledge challenges and constraints that exist in other cultures.

So, does the implementation of PM leads to better outcome in the public sector particularly in developing countries? According to Gao (2015), available evidence suggest mixed results where reform measures have led to improvement in some cases and failed on another. For instance, reflecting on their study on the implementation of PMSs in the Fijian public housing sector, Nath and Sharma (2014) conclude that changes in PMS is closer to myth as they did not improve operational efficiency. According to them, while international donor agencies have been practicing some influence on governments of the developing world to implement performance management systems, there has been some limitations to induce major changes and integrate these systems with day-to-day activities in the public sector. Another example is introduced by the study of Ohemeng (2009) on state owned enterprises in Ghana. Author concludes that the success of performance management will continue to be limited in developing countries due to number of institutional and capacity constraints that include limiting culture, institutional fragmentation, public apathy, and lack of leadership support. On the other hand, the study of Verheijen and Dobrolyubova (2007) point to the idea that the introduction of performance-based public management reforms in the public sector in developing countries can be successful if certain conditions are met. Building on their analysis of reform process in two Baltic States and Russia, authors make a bold statement by indicating that their findings contradicts the 'widely held notion' that performance management systems are not suitable for developing countries. Authors conclude that an incremental and step-by-step approach to reform along with sufficient political support and dedicated team for reform would lead to significant improvement in the efficiency and effectiveness of the public sector. This in their view will help to overcome systemic weaknesses that typify public organizations in that context. Mimba, Helden and Tillema (2007) and Tillema, Mimba and Helden (2010) refers to some of the characteristics of public organizations in developing countries which in their view could lead to unbalanced position between demand and supply of measures. Authors explain that reforms in the public sector are likely to lead to an increase demand for and supply of performance information only to be faced by some limitations such as low institutional capacity, limited involvement of stakeholders, high level of corruption, and high level of informality. Authors also acknowledge the

difficulty of generalizing their findings to all organizations in developing world by pointing that their line of reasoning does not capture all aspects and relationships in the public sector and call for more nuance and elaboration based on empirical works.

Research Methods

Introduction

This study is exploratory in nature based on descriptive and inductive qualitative research strategy. The study worked to examine the patterns of usage of performance measurement within a certain context- the Yemeni public sector- where there is little knowledge to determine specific construct, variables and relationships to be tested (Partington, 2002).

Two issues should be noted when it comes to qualitative research. First, as researchers get involved in data collection and analysis it should be expected that there might be some degree of subjectivity in terms of data interpretation. The subjectivity of data interpretation might indicate that if a similar research is conducted by other researchers they might come up with different conclusion. It is to that end why some researches need to be based on the work of many researchers instead of only one. Secondly and perhaps consequently, as with any qualitative research generalization is a challenge, and replicating the studies and coming up with the same findings and conclusions remains difficult in qualitative research. However, qualitative research is useful particularly when there is little knowledge about variables and theoretical foundation (Torres, Pina and Yetano, 2011). The conditionality of time and resources as indicated by Mikkelsen (2005) is another factor that could influence the research design by focusing on using people own words and interpretation to describe the level of significance of the issues at hand. Additionally, given the difficulty to monitor quantifiable performance indicators in certain contexts, qualitative approaches that are based on collecting data through extensive discussions with key stakeholders have better potential to enhance our understanding about performance improvement in the public sector (Sotirakou and Zeppou, 2006).

Purposeful Sampling versus Randomness

Given the political atmosphere at the time of interviews and based on discussion with my supervisor, I decided to adopt convenience and snow ball sampling. Convenience sampling gives the researcher more flexibility- in terms of time and cost- to determine the sources of data based on their relevance to the questions under investigation (Merkens, 2004). Selecting participants randomly would be helpful to make the sample representative and makes generalization about findings possible. However, given the type of conditions that apply to ethnographic studies, samples of this sort are often taken purposefully and not according to the principle of randomness (2004). Moreover, as with many qualitative studies, this study is less concerned with issues of

representativeness, which is less important in qualitative researches (Bryman and Bell, 2007), than understanding the area of the study itself.

Snowball sampling is a form of non-probability sampling that begins by identifying an appropriate respondent- one who is willing and can provide some answers to the research question(s). Upon the completion of the interview the researcher asks informants to identify another potential respondents who may be willing to participate whether in or outside his/her department or organization. The process continues until no more contacts are suggested by respondents or the researcher reach a level of saturation in which the new data do not seem to add much compared with the old one. According to Oliver (2006), the advantage of snowball sampling is that it demonstrate a sensitivity to potential participants with possible similar experience. It is dependent upon each participant understanding the nature of the study which allows them to nominate others who can add value to the research when it would otherwise be more difficult to achieve.

However, I'm also mindful of the limitations of this approach. For instance, the very fact that respondents acquainted with each other is a source of potential bias. Since snowball sampling is dependent upon nominations it is possible that next participants may not be necessarily suitable for the purpose of the research. They might have little information or no knowledge about the research issue or have no intention to participate in the research. Another limitation is that respondents could have biased understanding of the research issue or may share common features which make population homogenous which, in turn, undermines the generalizability of findings. Bearing that in mind, I worked to address these limitations by setting certain criteria for potential respondents as shall be explained in the following section.

Selecting the Sample

Respondents were selected based on their knowledge and experience in managing their own organizations. My interviews were limited to middle and upper level managers in public organizations in the Republic of Yemen.

Table 16 Public organizations by their types

Organizations	Type	No. Organizations
State's Ministries	Central Agency	5
Service Departments	Line Agency	4
Public University		1
Public Hospital		1
Public Fund		1
Public Service Company	State Enterprise	1
Total		13

Following Merkens (2004) guideline, I sought to have the following criteria: interviewees should have knowledge and experience in the research area, be articulate and capable of reflection, have time to be interviewed, and willing to take part in the investigation. All meetings were conducted in Arabic language and only 70% of interviews were audio taped while the views of the remaining 30% were included by taking notes during the interviews because some interviewees express their discontent for recording their meetings. Each interview would take a little more than 60 minutes on average with some exceptions in relation to some of the interviews that were taken in the second phase of data collection (45-60 minutes) due to the timing of the national dialogue conference's sessions.

Data Collection Stages

The first phase of the research involved live interviews with senior and middle level managers in five public organizations. It involved 12 semi-structured interviews with senior and middle level managers in five public organizations. All interviews were conducted at the interviewees working place upon their request. Asking interviewees to allocate part of their time for the interviews and visiting them at their organizations was not easy due to security issues and the difficulty of access to certain areas in the capital at the time of interviews- it might be helpful to note that some public agencies had to

change their location following the events of 2011 because of arms' conflict that created multiple confrontation zones.

The second stage of data collection was in line with an opportunity that presented itself in form of the National Dialogue Conference (NDC). Upon discussion with my research panel, I decided to use the forum for data collection, so I got acquainted with several members of the NDC including representatives of the steering committee. By presenting the issue to them they were happy to help and coordinate meetings with other members of the conference. Most interviews were facilitated and called upon by the steering committee based on the availability and willingness of members of the NDC to be interviewed during the event. This means, I had no influence or choice on the selection of participants except by indicating the abovementioned criteria. The interviews took place in a suitable corner in the Movenpick Sana'a Hotel- the NDC venue.

Based on the interview protocol, I would begin by illustrating the research issue, ask for permission to record the interview, and learn whether participants would like to remain anonymous.

Table 17 No. Interviewees by their title and type of organization

Organizations	No. interviewees	No. Interviewees by Job Position
State's Ministries	10	2 Ministers, 1 Vice minister, 4 deputies, 1 General director, 1 deputy director, 1 Minister's senior advisor
Service Agencies	8	2 Chairmen, 3 deputy Chairmen, 3 Senior Managers
Public University	3	1 Vice president, 1 Secretary General, 1 Dean of faculty
Public Hospital	2	2 Departments' Directors
Public Fund	3	3 Departments' Directors
Public Service Company	2	1 General Director, 1 Department Director
Total	28	

The third round of data collection lasted three days, when I asked for permission to make a presentation about the topic to members of the NDC. The presentation titled 'PMS in the public sector' aimed to give more information about the purpose of performance measurement and management and to take the views of a larger audience. Consequently, I was able to give my presentation to two main committees: the 'State Building Committee' and 'Comprehensive Development Committee'. Both committees had more than 120 members in total, many of which were senior public managers and officials. Following each presentation, the floor was opened for discussion and questions. The discussion that followed took the form of focus group meetings during

which I played the role of mediator, asking questions and encouraging the debate. It took an average of one hour and a half for each session.

Conducting the interviews

A semi structured interviews were developed with a pre-defined interview schedule Appendix 1. Taking in consideration that this is an exploratory project, the interview guide contained a fixed number of broad questions that aim to get a better insight about the research issue at the same time allowing the researcher to raise issues and probe further with additional questions. According to Yin (1994), an investigator should be completely prepared for the process of his/her study including the formation of good questions and possessing the skills to interpret the answers. It should be noted that while I worked to follow the interview schedule- that is to ask questions in the same order- that was not possible at all times. While the interview protocol included predefined open-ended questions, the specific order and phrasing of these questions were largely dependent on the context of each interview according to interviewee's interaction and retort with the questions. The course of some interviews might differ according to the situation. Bearing in mind the difficulty to get these interviews in the first place, the seniority of some respondents, and the sensitivity of the situation in the country following the Arab Spring 2011, the focus was to get as much information as possible and to allow respondents to express themselves as freely as possible. Bearing in mind that the researcher must be able to interact quickly with the flow of answers and pose new questions when necessary (Gray, 2004), I encouraged interviewees to provide further explanations about their views and probed with additional questions when deemed necessary or when an interesting topic came up.

Data Analysis

The analysis process started prior to the transcription of the audio material where I tend to listen to each recorded interviews for the purpose of taking notes and plan for next round of interviews. This early stage of notes-taking was helpful for later stages of data analysis as it enabled me to compare between old notes and new ones that were taken months after the data collection stage was over.

Audio recorded interviews and notes were then carefully transcribed and checked numerously to avoid possible data loss and to ensure data integrity.

The analytical strategy followed the approach suggested by Miles and Huberman (1994) who proposed certain stages of data analysis which can be summarized in the following steps:

- The first stage of the analysis started by carefully examining transcribed material more than once in order to be familiar with the data and to determine possible

errors or data loss during transcription. Notes were also taken to identify general themes and possible categories that could be helpful for coding in a later stage.

- Transcribed material and field notes were then imported into NVivo for coding. NVivo is a qualitative data analysis application that enables users to organize, analyze and visualize information. The software automatically keeps track of information and facilitates direct referencing and nodding. Selected data would be marked by identifying relevant category, code, and page and line number from the original transcripts. NVivo represents the analytical categories as nodes and data as sources.
- A summary document form was created for each interview which contains a brief summary of interviewees, time, and location of the interviews. According to (Miles and Huberman 1994) this kind of documents provides a useful information about the context of each interview which permit rapid retrieval of important information and can be coded for later analysis. This was achieved by linking each interview to a distinct document that contains the above-mentioned information through the use of a special function in NVivo called 'Memo Link'.
- After all that the coding process began. In qualitative research, coding refers to the process by which data are broken down into component parts with particular names (Bryman and Bell, 2007) or simply relating particular passages to one category (Schmidt, 2004). Initial coding started by reviewing data to identify possible themes and general categories. This was achieved by repeated reading and intensive scrutiny of each transcript.
- Consequently, a more thorough process of review and refinement of textual data were conducted to assess and classify important and relevant parts. Transcribed material was first coded into free nodes. As outlined by Miles and Huberman (1994), when data are reviewed line by line for labeling and categorizing a list of provisional codes get created and continue to grow. However, it should be noted that while the aim was to identify possible patterns on data I was also mindful to note different topics and individual aspects that could be related to the research problem. According to Berg (1998), the main purpose of open coding is to open inquiry widely. While possible interpretation, questions, and some answers could emerge during this stage it is important to bear in mind that nothing is conclusive until the whole process of coding is over.
- As free coding continues, NVivo automatically generates a tree of all nodes which lists all the created nodes during the process. Miles and Huberman (1994) indicate that in order to be able to compare the cases with regard to dominant tendencies, the quantity of information has to be reduced at some point of the analysis procedure. Consequently, the second stage of coding started by re-examining and splitting the tree-nodes into shorter and more meaningful version.

This entails re-arranging nodes according to the main themes and the creation of parents and child-nodes which are related to broader topics or categories.

- As the coding process is about to conclude I went on to reflect on the data and emergent themes in relation to the literature. The aim was not merely to unveil possible themes that was discovered by the literature and could be missed by the analysis but rather to compare findings with that in the literature to discover similarities or dissimilarities.
- The final level of analysis was handled through the arrangement of parents and child nodes into headings for the write up. By showing each respondent's response in relation to relevant themes it was possible to draw conclusion. Chapter headings were based on a general summary list of main findings that was created to represent answers for all the research questions along with any emergent and additional important findings. The list was then linked to and backed up with quotes and selected text from the final tree-nodes.

Translation Issues

All the interviews were conducted in the Arabic language, the native tongue of both the interviewer and interviewees. Similarly, and in order to maintain the integrity of data and avoid any possible loss of meaning the analysis was conducted in the same language. That means that all transcribed material was imported into NVivo in the Arabic format. The process was not as straight forward as it is for English material since the software isn't fully compatible with the Arabic language. After several attempts and some help from friends I was able to overcome this issue by importing the material into NVivo in the document format after changing the type of font in those documents. It was not completely flexible though, but it seems to be sufficient for the purpose of analysis. Further, only general conclusion and findings were then documented and taken in English. The entire process was conducted with the outmost attention to be careful in conveying the meaning and the outcome of the analysis.

The Yemeni Context

Overview

Describing the public sector in Yemen based on the literature seems a tiring task because of limited empirical and scholarly research on the country. Available documents and international agencies' reports such as World Bank, IMF, UNDP, USAID, etc., might be helpful in some aspects but they do not provide a full picture about performance management in the Yemeni public sector- as they mostly focus on strategies and major programs such as poverty reduction, food security, and economic growth and development. However, a quick review to some of these reports supports the notion that weak capacity of governance institutions has been considered one of the main challenges to economic performance and growth. The quality of governance continues to deteriorate and scores low in the World Bank measure of governance quality and the government have often been accused of inefficiency, too much bureaucracy, corruption, and negative political influence that delays project implementation (World Bank, 2008a; OECD, 2010). Problems of the Yemeni public sector can be explained further by analyzing political and socioeconomic conditions in the country.

Political and Economic Challenges

Yemen is a low-income and a fragile state that is faced by many political, economic and social challenges and problems. It is one of the poorest countries in the world with less than 600\$ per capita GDP (World Bank, 2008b). With highest population growth in the world, more than third of its people is poor and the country ranked low at the human development index report (2008). Nearly half of its citizens are young below the age of 20 and the country has limited resources and simple infrastructure, and with severe shortage in water and high fragmentation of people (Moriani, Al-Hammadi and Al-Zawm, 2013; OECD, 2010).

Yemen has faced many political, economic, and social challenges over the years. According to the World Bank reports, following the year 1990 Yemen needed to deal with the aftermath of unification that include but not limited to fiscal imbalances, increase of inflation rates, and a decline in total revenues as opposed to an increase in total spending. The size of the civil service has more than doubled and government's efforts to control spending within the resources' limits were not successful which explain why the country has always needed foreign assistance. During the same period more than 800,000 of Yemeni workers were deported from Gulf countries in a move to punish Yemen on its stance against the Gulf War depriving the country from immigrants' remittance and financial aid that used to support its economy (World Bank, 2008b).

Political tensions amplified following reunification in 1990 between political rivals for four years, the interim period, just to end with a civil war that lasted for two months and

consumed considerable resources. Budget deficit increased over the years following the civil war due to higher spending particularly on wages and on infrastructure particularly in the south. Macro-economic challenges continue to increase due to additional problems the country had to face over the following years such as separatist movement in the south, six wars against insurgency in the northern part of Sada'a, flood of thousands of refugees from the horn of Africa, tribal kidnapping of tourists, and active terrorist networks. Further, since the country depends on oil export for more than 75% of its budget a declining oil reserves deprived the country from an important resource (World Bank, 2008a) (IMF, 2013).

In 2011 mass demonstrations in the streets of Tunisia that resulted in the oust of President Bin Ali and similar scenes in other Arab countries on what would be called later the 'Arab Spring' inspired protesters in Yemen calling for a regime change. The events took another twist in Yemen where what started as peaceful demonstrations turned into a political crisis that witnessed some brutal clashes and an attempt to assassinate the president. After one year of political stagnation and tension the crisis seemed to be finally resolved by political agreement between political rivals, brokered by countries of the Gulf Cooperation Council (GCC) and other international partners of Yemen. The signing of the GCC Initiative agreement was thought to end the political crises and saved Yemen from a terrible civil war. The transition included the resignation of the former president, the composition of a National Dialogue Conference (NDC), the formation of a national reconciliation government headed by the opposition, and the election of the vice president as the new head of state in an election that had no other candidate. The NDC aimed at discussing and agreeing on a road map for the future to resolve long term problems of Yemen. The NDC concluded in January 2014 with high expectations from the outside world but with great level of cautions from national political parties who continued to show their resentment of the way the NDC was managed by the new president.

With time passing the general public perception is that nothing has changed under the new leadership and that current transition government was not much different from the former if not worst. Fiscal challenges continued to increase as attacks on gas and oil pipelines continued and deliberate sabotages of public properties and services such as electricity towers escalated in addition to the inability of the transitional government to generate additional and urgently needed resources. Given the severe economic and political conditions in 2014, and in order to generate enough resources for government programs many international partners advised or rather pushed the government to choose between devaluating currency or cutting subsidies on diesel and gasoline. When the transitional government failed to make up its mind on the issue and under fear to lose foreign support, president Hadi announced a bold decision to fully cut off subsidies and liberalize fuel prices. The wide public discontent and anger was used by Al Hothies- a political group that was represented in the transitional government and part of the opposition against the former regime- to advance towards the capital through demonstration and organized sit-ins in different parts of the capital Sana'a. The events

following that were dramatic and lift the country in a direction that is marked by a serious, expanded and devastating civil war.

Reform in the Public Sector

In line with economic difficulties and conditionality of donors, there have been several initiatives to reform the public sector in Yemen. For example, in 1998 with the support of donors, notably the World Bank, the government approved a strategic framework for modernizing civil service and financial management. The main focus of reforms was essentially to decompress civil service salary to the minimum in order to save resources for other pressing issues. To understand that, one should first know that following unification in 1990 the civil service in Yemen was not only underpaid but also overstaffed as a result of merging public institutions in the former two states. More than two thirds of current expenditure goes for wages and subsidies which according to experts one of the highest levels in the region (IMF, 2013). According to the World Bank (2008a), the reform objectives include:

- Streamlining government and eliminate duplication and non-essential services;
- Increasing the quantity, quality, and cost effectiveness of public services;
- Increase transparency of public procedures
- Improve financial management systems to enhance transparency and combat corruption
- Restructuring public employment system
- Developing personnel management system

As it turns out, reform program was not successful and it was not possible to achieve all objectives. For instance, one of the outcomes of reform was the establishment of anti-corruption commission with free access to financial records and performance reports of all public entities. Yet, in 2008 the World Bank in its anti-corruption reforms reports claim that country financial accountability assessment suggested progress was limited and the commission itself needed more transparency and lacked proper support from the government. Further, while there has been reduction on some illegal payroll and formation of salary structure based on merit, government efforts to reduce the size of the civil service was not only limited but have been described as an utter failure (World Bank, 2008a).

As a result, the Yemeni government with the support and coordination with many of its international development partners declared an ambitious reform program named the National Reform Agenda (NRA) aimed at strengthening governance and improve investment climate and democratic institutions in the country. The key components of the NRA were: decentralizing services, goods and public works; modernizing public procurement; and combating corruption. The agenda for reform meant to cover many

areas such as legal measures to combat corruption, improving the judicial system, political and democratic development, advancing and redefining the civil service, and promoting economy and enabling business policies. The following are some of what thought to be some of the achievements of the NRA:

- Launch of the National Anti-Corruption Campaign
- Ratification of the Financial Disclosure Law
- Approval of procurement manual and standard bidding documents
- The adoption of the Anti-corruption Law
- Government's acquiescence to the Extractive Industries Transparency Initiative
- Adoption of the Public Finance Management Reform Strategy PFMRS
- The establishment of the Supreme National Authority for Combating Corruption SBACC and High Tender Board HTB
- Adoption of new Tender Law

Yet, there are signs that discontinuity caused by political tensions and economic difficulties along with the lack of political commitment remains serious challenges for a complete reform in Yemen (OECD, 2010). Capacity of government to act and to take serious decisions for reforms remains an issue for question with international partners who publicly doubted government's ability to effectively and transparently spend large amount of donors' money for development programs on time. Referring to the case of Yemen an independent evaluation group by the World Bank (2008b) concluded that government may have adopted reform strategies and passed new legislations but when it comes to implementation there is slow in momentum, delays in action, and even, at some points, a complete halt of projects or certain program. One lesson that some donors noticed was that the support of few officials shouldn't be mistaken with full-fledged government support.

Findings and Analysis

Before proceeding to present findings, it should be noted that while understanding the key patterns of PM usage in the Yemeni public sector should reflect general trend, it should also account for exceptions. This is necessary because I noticed that speaking about the public sector as one might be misleading given the differences that exist between different types of public organizations. The Yemeni public sector is large and complex consisting of various types of institutions that varies significantly in their size and types of services. Similar to other countries, most of public organizations including the central government are based in the capital with different branches in different provinces of the country.

Bearing that in mind, I decided to follow Ohemeng (2009) in dividing the public sector in Yemen into three main categories: (a) central agencies (b) line agencies and (c) state-owned enterprises. Central agencies refer to public institutions that are responsible for policy management and control, priority setting and coordination, and planning and managing financial and human resources. They have formal and informal authority over other government agencies and often direct their actions. Line agencies, on the other hand, are organizations that report to upper government levels, fall under their control and guidance, and relay on government support for all or part of their salaries and operational expenses. While these organizations known to provide direct services to citizens they normally have no mandate or authority to direct other agencies or their operations. State-owned enterprises are public entities that although owned by the government, have closer link to private companies in terms of operation, control, self-finance, and focus on financial gains. Both similarities and differences between these organizations will be discussed in the course of this section whenever possible.

Understanding the concept

Several questions were asked to examine the perception of interviewees about measurement and its use. One way of gaining an understanding of their experience and familiarity of measurement was to ask informants about how they manage performance and how they would know if their organizations performed well or not. The following are some issues that were deduced from their answers.

All interviewees indicate that performance is related to pre-set objectives that are usually defined in the annual plan. Each public organization would have to develop an annual plan that set out strategic goals and major activities. Determining efficiency and quality of performance is dependent on the extent of an organization ability to execute its plan. A typical answer would look like this one:

We develop a plan for our work by the beginning of each year and by the end of the year we prepare a final report for what has been done over that year. Then, we make a

comparison between what was planned and executed in order to determine our achievement.

However, while annual plans expected to reflect main vision, responsibilities, key strategies, and main targets, our interviews reveal that annual planning in most of the organizations covered by this study, mostly central and line agencies is rather normal routine than being a real tool for managing performance. For example, respondent's answers reveal that annual plans of most central organizations contains broad goals only with no specific or measurable targets, no clear description of policies or activities to achieve goals, and no performance expectations. Probing questions like: what are your outputs, products or services? What is your desired outcome? Who are the customers? All seem to be daunting for many respondents to answer. It shows that there seem to be a problem in goal identification and planning which negatively impact their ability to determine organizational performance.

Further, most interviewees seem less familiar with the concept of measurement and unable to give examples of use of performance measures or indicators. While they all agree that measuring performance is 'good' or 'important', the majority of interviewees admit that there has been no specific guidance, training, or prior experience in measuring performance. For example, almost all respondents used the term 'performance evaluation' at some point or another to refer to the measurement process or to the benefits of measurement. This reflects a lack of understanding about the purpose of measurement and that they perceive measurement as a synonym of evaluation.

Performance measurement is essential. Without evaluation of performance it is difficult to determine what needs to be done to improve.

Performance measurement is important for evaluation if it's not evaluation itself.

If there is no success there is failure and we must find out the causes of problems and work to fix them. That is why it is important to evaluate performance from time to time and compare it to our annual plan... I'm not sure though that we are measuring performance in the right way.

Few interviewees mentioned other purposes for measurement, but even those mentioned came in line with evaluation or as a result of evaluation.

Without evaluation there are no better decisions and planning and consequently there is no improvement in performance.

Performance is measured in relation to predefined goals for the purpose of evaluation. We have many departments, each with certain plans and goals and responsible for reporting their performance and the reasons for success or failure.

Further, it was a bit surprising to see several respondents refer to appraisal when speaking about organizational performance which imply their inability to see the difference. These interviewees suggest that a formal assessment report is issued annually for each employee to determine their performance and commitment to their jobs. It is a top-down process that is based on subjective evaluation of workers by their managers or by the human resources' department. Based on their view, appraisal could reflect organizational performance in most public organizations.

We used to have a performance review by the director general; He would assess each one of us. But not anymore, even incentives are not linked to the evaluation because no one gets excellent everyone gets very good rating.

I consider it a weakness in the work of the organization and its leadership. Because previous administration was keen to evaluate; and people were convinced that they receive financial incentive based on their evaluation.

We use some measures for determining efficiency and effectiveness for our performance. For instance, groups' leaders would have to review employee's behavior, commitment, appearance, development, and work performance.

At some point, when I explained that the question was about organizational performance one respondent replayed by saying:

Employees' collective performance would reflect on the organizational performance on general...it means that there is really a relationship one way or another...in the west you might invalidate the practice license from a person which makes him careful, exerting effort and enhance the quality.

Few others would confuse organizational performance with major indicators about certain sectors in the state:

... The real measures have been developed by international organizations such as WHO and OXFAM which visit Yemen. They are the ones which conduct research and if you go back to their statistics you would find the country in the bottom in all these measures.

Further, some interviews suggest that annual plan should be realistic with little space for change or developing new goals due to many constraints, mainly financial. Hence, rather than spending some time thinking about objectives and targets for the upcoming year and consequently thinking and discussing means to achieve them, managers would have an expectation about the day to day business and the final outcome from 'doing business as usual' and define their objectives accordingly.

I think that the difficulties lie on the fact that plans and objectives are not based on what is there in reality...., many of those plans and objectives were not realistic but rather exaggerated.

... We would have predetermined goals but according to a flexible and not confined framework because we make our assessment but consider other elements that might be beyond our control, possibly under the control of other entities such as Ministries of Finance, Civil Service and Planning

Development of annual plans and goals could take either the participatory form where lower department prepare their part and send it to upper levels or be the responsibility of few managers or special department. A bottom up approach seems to be dominant on organizations with clearer type of services or outputs and the process is more central in state ministries where it is limited to either few senior managers or a special department for planning and evaluation.

Source and Types of PI Demand and Supply

When respondents were asked how they share performance data they all seem to suggest that both internal and external stakeholders would have to depend on annual or semi-annual reports. They indicate that as part of public reforms initiatives public organizations are required to report to higher level authority about their activities. Central agencies like ministries would submit their reports to the office of cabinet and the office of the head of the state, while other type of public organizations would report to central authorities like relevant ministries or financing agencies. Public managers explain that performance is related to pre-set objectives that are usually defined in the annual plan. A public organization is expected to issue two types of reports each year: financial closure statement and operational performance report. Preparing these reports should begin by the end of the fiscal year and ready for distribution during the first three months of the next year. Budget statements are prepared by government agencies each year with limited details about organizational outputs and possible outcomes of financed programs. Annual performance reports represent an assessment of government work and organization's ability to execute their plans and to perform according to their objectives.

Findings suggest that performance review in most public organizations is mostly qualitative rather than quantitative. It is generally based on subjective evaluation by senior and upper-level management and their personal assessment of the work implemented during the year. Such assessment is not systematic, i.e. based on certain and orderly procedures, but rather on management opinion about the overall performance. Moreover, findings also show that for the majority of government agencies there are simple but not sufficient guidelines that provide clear information on the development of performance reports. Many interviewees indicate that their organizations did not receive training in the way that performance should be reported or measured.

Further, findings also show that demand for performance information is relatively weak in the Yemeni public sector. According to interviewees, there is no real demand for performance information more than it is an obligation for reporting performance. As outlined by one interviewee:

...apart from annual reports we aren't normally asked for further performance information.

Some interviewees provide plausible reasons by indicating that performance information is rarely used for performance improvement or accountability purposes. They indicate that except for few regular check-ups by Ministry of Finance, mostly on financial records, there is hardly any request for performance data and there is little evidence of an oversight body holding a public agency or managers accountable based on performance data. According to them, critical functions such as planning, budgeting, or even the appointment of senior officials are not necessarily dependent on information pertaining to performance. Interviewees indicate that even when they report their performance to upper levels there is hardly any feedback on their reports. The following quotes illustrate how interviewees elaborate extensively about this issue.

It is really up to the leadership such as the minister. If he is interested he will show his interest and if not he might only read it with no comment and sometimes he wouldn't even bother to skim through.

'we have the cabinet office that request two things our performance report and annual plan for next year which should help in monitoring and evaluating our activities....., but I hope someone would discuss that with me or tell us you have succeed in doing this or failed on that, that doesn't exist.'

We only report to them but there is no feedback which is wrong. They should have some comments on our reports

I think our reports might be of interest when something happens. For instance, if the media tackled that negatively or there is a general dissatisfaction of our services they might come to question the reasons and look into our reports, but we rarely receive any feedback.

Despite our efforts we do not get proper feedback. There is no evaluation of our reports. They take these reports, print them, and returned to us as a published material, but no one reviews our evaluation

Furthermore, findings support the notion that pressure from stakeholders has an influence on the production of specific performance information. Data indicate that the supply of performance information correspond to the source and type of demand. It shows that demand from external stakeholders most notably funding bodies do have an influence on the way that financial and non-financial performance should be reported.

For example, interviewees indicate that there is a certain format, designed by the Ministry of Finance, for reporting expenses in all public organizations. While some expressed their dissatisfaction about the format they all seem to agree that, to receive money from the government, all public organization should adhere to predetermined roles and arrangements. Similarly, our interviews show that public organizations that receive financial support from donors need to utilize some sort of measures or indicators to comply with external requirements for monitoring financial and non-financial performance. Several interviewees pointed to the fact that external donors have some standards for monitoring and controlling their support which impact the reporting process.

We need to consider the opinion of those who finance us when designing our programs to ensure their support; sometimes we seek to explore their input about a particular field, design of a certain project, and the way we should report progress

'We do our best to accommodate their views and requests but without compromising ours of course'.

Worth to mention that findings from this study suggest that pressure from specific stakeholders such as donors is not always positive. Managers explained that meeting expectations of funding agencies is not easy and creates many challenges. Their views reflect a possibility of conflict of interest between internal and external stakeholders over priorities and working process.

'Their sole interest is on the financials records. They want to say that they utilized the resources efficiently. They also request the use of some special indicators to measure the efficiency of their support; this means more work and only focused on the output and not the outcome'

'sometimes there are problems like when the Europeans insisted on following their approach on reporting performance and money transfer; we used to do it differently with other donors like UN organizations, so we weren't trained for the European style; it took us two years to get the money for one program, but it is ok now; we changed our approach to match their needs'.

Findings indicate that performance information is not usually published or disclosed to the public and that there is limited role for citizens and supervisory entities in relation to the demand for and consumption of that information. For instance, while statutory and supervisory entities might have an access to performance information they seem to have little influence on the way performance should be reported. Interviewees indicated that it is not common for legislators to ask for performance data or to provide feedback or guideline on performance reports. This suggest that performance information is limited to the use of more powerful stakeholders mainly managers and financing bodies.

Furthermore, most public organizations do not engage in discussing their performance with other entities outside their authority. In fact some managers expressed some concerns over the possibility of publishing performance reports and gave interesting thoughts about their views:

' it will increase public discontent; I mean given current bad public performance, these reports, if transparent, will represent an admission of failure and will be used for pointing fingers towards the administration- any administration.

Don't get me wrong, I fully support transparency and sharing information with people, but I doubt that the government would be keen to do it.

Consequently, interviewees also suggested that there are limited efforts for measuring citizens' satisfaction about public services. As indicated above, even when public opinion seem to matter for some managers, measuring the level of satisfaction of citizens seem to be dependent on subjective and personal judgment of managers based on their own interaction with people- whether they hear compliments or criticism of their work. Two managers provided what might be considered an explanation for not incorporating specific measures for customer satisfaction:

'It might be difficult to evaluate the satisfaction of people; we deal with different types of people in different places and we need a lot of time and resources to survey their satisfaction'

' we don't need to evaluate; when people are not pleased from our services they let us know; they complain about it directly or through the media; and they always complain regardless how hard we work to improve, they are never happy'.

Main Barriers

Findings indicate that there are many other limitations to the development and use of measurement in public organizations in Yemen. The following are the main factors cited by interviewees.

Lack of financial resources

Most prevalent comments related to the lack of proper financial resources for public work. Most managers stressed on this issue and explained that there is inadequate policy for allocation of resources according to the needs of their work. According to them, this directly influence performance and undermine any attempt for reform in the public sector. Interviewees expressed their understanding about fiscal constraints and financial limitations but they also were keen to reflect their discontent about the way priorities were managed when it comes to budgeting. According to some, the process of

allocating budgets for public entities is not rational- based on organizational goals- but rather dictated by Ministry of Finance or Ministry of Planning.

'...there are very limited resources in the country and that impact us because we get less money than needed to attain our objectives'.

'Financial resources is an issue; When something happens and affects our finance we can't implement certain programs... at any point of time during the year we might need to adjust original plans to accommodate with change in our budget...'

Managers also pointed out that financial limitation is not limited to budgeting only but also related to other problems such as the difficulties of getting money on time and the bureaucracy of settling their accounts with the Ministry of Finance or getting the Ministry's approval for financial statements.

These financial aspects are very complicated because we are bound by the regulations of the Ministry of Finance. Imagine when you already know that a program would cost ten million only to find that they approve five or three million. How would you measure your performance then if you even are not able to attain your goals?!

It is unfortunate. the budget is determined by the ministry of Finance but to get the money and spend it we need to work it out in pieces where you can't get one payment until you settle the first; that is a lot of paper work. Even after that nothing is certain, only salaries and maybe utilities get to be paid with less hassle.

The government would have to allocate resources according to its priorities, but I doubt it does it correctly.

The exception was public organizations that were less dependent on government's support. Interviewees in this group are working in organizations that enjoy some autonomy in management and have other means of finance- either by collecting direct taxes or service charges, or by receiving money from external sources such as donors. Managers in this group were less concerned about the availability of money more than they are about justification of expenses.

Financial resources are not an issue here. You can say that we are more concerned about how we spend it.

The money provided by the government is mostly for operating expenses such as salaries and utilities, but as for programs that we work on, we often present them to donors.

Limited skills and qualifications

The second most noted barrier by interviewees is inadequate human resources in public organizations. Managers explain that public organizations in Yemen are known to have a chronicle problem of inadequate skills and qualifications. The following comment

presents an indication of the level of dissatisfaction about the qualification of public workers:

‘if you ask me how many of them [employees] are efficient, those you feel perform well and that the organization is dependent on their work, I would say less than 40% of the total labor force’.

‘People are the base; humans make success or failure. If you have strong cadre you can achieve whatever you want and without it you achieve nothing, and this is our main problem human resources’.

‘one barrier is the qualifications of people; having the wrong person in the wrong place’.

Some respondents reflected on the difficulty to obtain talents and skilful labor to the public sector. Their views indicate that public work is not attractive for talented people for many reasons that include but not limited to: low payment and limited incentives compared to that in the private sector; lack of training culture in the majority of public organizations; hiring could be politically motivated and not based on merit or specialty; overstaffing or inflation on the number of employees in public work.

Nevertheless, it should be noted that managers in non-central agencies were less complaining about qualifications. A number of those managers pointed that the reason they have better labor force than other organizations in the public sector is related to the flexibility given to them- granted to them institutionally- to hire employees directly which helped these organizations to get people with certain skills that correspond to their requirements. Interviewees pointed to the fact that a centralized policy for hiring public workers in which the Ministry of Civil Service, not government agencies themselves, have the authority to announce vacancies and select candidates for a particular job seem to be one important reason for the increase of inadequate labor in the public sector.

Resistance and Cultural Barrier

Many interviewees pointed to the cultural barrier for the development and use of measurement in public organizations. Whenever the issue of culture was raised, interviewees were encouraged to elaborate and provide further explanation about what they mean by ‘culture’. Most of the answers suggest that principles of measurement might conflict with the predominant working practice which creates resistance. Interviewees explained that the introduction of a reform tool such as measurement expected to face a great deal of resistance and little acceptance by managers and workers alike. Their answers revealed that performance information is related to the overall evaluation of the annual plan which is mostly based on subjective and qualitative assessment rather than actual measurement or use of indicators. This has been the practice in the majority of public organizations for a long time and hence has become part of their culture. One manager, who is familiar with the function of

measurement, indicates that changing or challenging an existing practice is both difficult and time consuming:

‘People are used to a particular way for determining performance... introducing a new system might not appeal for both people who measure and those subject to measurement’.

The analysis signifies that the dominant culture of public organization is a culture of leniency and ease when it comes to an assessment of public work. For instance, managers pointed out that since performance information are not often used to take decisions or actions, holding people accountable for their actions or performance is not part of their culture; *‘it is not that common’*.

Moreover, it was interesting to note some comments on the value of measurement in relation to existing culture. According to some managers, measurement might indicate more accountability but if it is not applied properly and no one gets accountable, it would eventually have no value.

‘People should feel that measurement has a value. I mean what is the use of measurement if it was to be put in drawers without extracting results and using them for action that focus on attaining objectives and improvement?’.

‘we need to be able to say well done to those who performed well and bad work to those who didn’t. this is so important in performance review... in a way that creates motivational contest.’

‘this might defeat the purpose of measurement and would be seen by people as useless tool’.

Several other managers pointed out that if the use of measurement would mean more accountability it would be used as a tool for punishment. They explained that performance will likely remain low in public organizations for many reasons, all beyond their control, notably because of shortages in finance. According to them, in order to mitigate resistance internal stakeholders need to know and believe that the new system will be used in their benefits and not against it.

‘Measurement systems need a culture too, because they wouldn’t be accepted from those who would be subject to these systems or even sometimes from those who apply them and this is a problem, exactly like democracy; therefore, measurement systems need to be linked with some sort of public awareness that it is not a tool for punishment, as some would look at it, but for track correction and for performance and work improvement which is the biggest challenge in my opinion’.

Lack of Incentives

Seven managers pointed to the lack of proper reward and incentive system in public organizations. According to them, within the context of the Yemeni public sector, it is difficult to get financial rewards in general and it is even harder to link organizational performance to the work of individuals. Managers indicate that there is a need to motivate workers and warn that without a clear link between performance and rewards public workers will be less receptive to measurement and more likely to resist it. However, managers also suggest that limited resources and bureaucratic restrictions undermine the possibility to reward or sanction good and bad performance respectively.

Employees are complaining lately from the lack of rewards and punishment system in our organization, but given that we have less money to spend we can't give financial rewards such as bonuses or increase in salary particularly to all employees regardless of their performance'.

'This is not a private company, we are simply bounded by money deficiency and rigid regulations imposed on us'.

There is no link between incentives and evaluation; no one get an excellent grade everyone gets very good and those who are absent or perform bad get a good or average grade.

The situation remains the same even for those organizations which have less difficulty in the resources issue. It was noted that even in those organizations there was no clear link between performance and pay. Managers justify it by saying that first *'regulations are rigid'* and second salaries are usually much higher in their organizations than other public organizations and so is their *'expectation of the work of employees'*. While some explained that they tend to use non-financial incentives to reward workers they did not provide clear examples about that.

Additionally, few interviewees pointed out that culture change and establishing a new practice will require a lot of time and need consistency which is not guaranteed in the public sector. One manager explained by saying: *'the longer any program or change takes the weaker it becomes'*. Another explained that as time passes without imposing change as well as reaping the benefits of change *'people will learn smart ways to bypass the new system... it is their way to resist and say no to progress'*.

Additional barriers

The findings of this study reveal that there are additional barriers to the development and use of measures in the Yemeni public sector which were not necessarily noted by interviewees directly. Five main factors could be identified.

The first factor is related to the view of interviewees about possibility to ‘determine’-measure- performance objectively. While interviews indicate that there is limited endeavor to measure cost-effectiveness or the outcome of particular program or intervention in the majority of public organizations in the Yemeni context, findings also suggest that the more clear objectives are the easier to measure progress to achieve these objectives. When interviewees were encouraged to give some examples about their goals, outputs, and outcomes there was a distinguishing difference in their answers. For example, I noted that interviewees in central agencies like ministries were having more difficulty to clearly define their outputs and expected outcomes. For the majority of interviewees in this classification, goals are often ambiguous or continue to shift according to political agenda. As such, organizational objectives seemed to reflect activity descriptions more than a definition of clear and measurable goals. Their responses indicate that linking their objectives to the overall strategy is a major challenge. Several interviewees admitted that they conduct their business without necessarily linking their activities to the overall objectives or strategies; others were keen to show their interest or their efforts to address this challenge:

‘We recognize that we have a problem. A special consultant is now helping in defining our objectives clearly and developing our strategy and plans for the coming period. There has been a deficiency but we are working on it’.

Understandably, respondents from line agencies which provide direct services to the public seemed more comfortable in terms of defining outputs. Managers in this category explained that they focus their attention in meeting their targets as outlined in the annual plan.

‘...our target is ten thousand researches; this is an indicator which we should accomplish by the end of the year... these targets are part of the overall objectives...we report progress to upper levels...there is no indicators for everything’.

The fact that interviewees who pointed to goals ambiguity raised the measurement question boosts the link between the two. However, the interviews show that some aspects of performance are beyond measurement which might explain the need for managers’ skills and expertise for evaluation purposes.

‘This car that you spoke about, if I use a simple Japanese car it is easy to know how it works, I could tell if fuel is enough and I could read other indicators; but if you bring me an electronic and sophisticated car with numbers, things might get complicated for me. Hence, measurement mechanism should be really proportionate with people’s capabilities to apply the principle of measurement’.*

‘How could we determine uncertain things? If we are not sure we are able to meet our objectives we certainly can’t speak about performance’.

Secondly, the lack of communication and information sharing or access to data in most of public organizations could be yet another limitation. Based on answers on the first

question, it seems that public organizations have long hierarchy of command that include too many levels and there is deficiency in collecting, transferring and communicating data to different levels inside and outside the organization. Some managers questioned the validity of current performance data and explained that most organizations still use narrative reports to report performance which is why the content of many of these reports cannot be trusted for decision purposes.

'there is a deficiency in managing information in government organizations; there is no interest in automation, no documentation, no standardization in documentation.

'Do you see constant forms for data collection in work? No you don't. There are no data or information forms that can be used in statistical reports'.

'performance review is dependent on the data, getting data is difficult. For example, I have the right to ask for some information from my colleagues but when I do they don't replay and no one seems to force them to cooperate'. There is a problem. If I apply measurement how could I use it if data flow is not right and organized?'

'there is no standard for performance review. It is subjective. How is it, a performance review would change from poor to excellent just because the head of the department has changed. This creates a problem of course, because reviewer is either, biased and sympathetic, or wasn't trained well on performance review'.

Thirdly, several interviewees have complained from external interferences from other stakeholders which in their opinion affect their ability to both plan and identify their objectives as well as measure their performance. This shows a power struggle between different stakeholders especially when it comes to resource distribution.

the ministry of Finance have the right to make changes in any budget even if it was already approved by the parliament..., for that, measuring performance would be a difficult task.

Some managers commented on the lack of coordination between internal and external stakeholders when it comes to measurement. According to their view, PM should have the same meaning for all and need to be used to maintain its value.

'Even if we applied measurement in our organization those in upper levels who follow and evaluate us is not interested; if you start feeding them equations and standard measures they would reject it; ... if upper authorities do not understand all of this, then what? It might be over their knowledge and administrative capabilities; it doesn't work this way, those who execute is not the one who impose measurement it should be the opposite'.

Finally, and perhaps most importantly, it was interesting to note that only eight managers referred to the lack of management support and leadership as a barrier for the development and use of measures. While this might indicate that senior public managers are supportive for change it could also reflect the opposite: that PM is not used

extensively in managing performance in the public sector or that public officials do not want to admit any wrongdoing from their part. Analyzing interviewees' responses to various probing questions seemed helpful to provide an overview of the relationship between political support and managing performance. For example, it seems that there is undefined responsibilities and lack of clear command and structure which limit managerial support. Interviewees' answers show that public officials seem to be occupied and overwhelmed with pressing issues- that is solving daily problems- rather than thinking of innovative changes and performance improvement. One interviewee described that by saying:

'it provides a sense of false impression of busyness and self-importance which block innovation and possibility for change'.

Further, two contrasting issues seem to matter in relation to internal managerial support. First, findings suggest that there is limited degree of discretion given to managers in the public sector who seem to be confined by many formal roles and bureaucracy. Their authority and ability to act based on performance data seems to be limited by lack of finance, rigid regulations, and undefined responsibilities. Second, findings suggest that public managers have the authority to present information in a way that maximize their interest and avoid unwanted attention. This means that internal management could shield itself from criticism by gaming performance data indicating that the desire for adopting measurement- that might expose bad performance- would rather be limited as well. Examining all responses to the first question suggest that managers are not experienced in measurement and that they have not been affected by the use of measures. Hence, managers might want to show only their openness to change and not their reluctance to and possibly fear from new approaches.

Discussion

This paper has examined the patterns of performance measurement and management and their application and use in the Yemeni public sector. It shows that the Yemeni public sector continues to deteriorate and suffer from chronic problems of mismanagement and lack of needed capabilities to perform its duties. The Yemeni experience shows that despite several initiatives for change in the public sector, performance improvement has not been a major concern in past administrative reforms. It reveals that changes in management control that took place in many western countries were not followed by the same in the non-western world.

The Yemeni case indicates that one of the main challenges for the utilization of measurement in the public sector within the context of developing countries is related to the lack of awareness and understanding of the use and usefulness of measurement. It shows that performance in the public sector is determined by classical methods that are based on subjective evaluation rather than actual measurement of performance. The study reveals that there are cognitive limitations that prevent public organizations from utilizing measurement by indicating that key actors are not familiar with the concept of measurement and its functions. This includes a lack of knowledge about the value of performance measurement, lack of skills to design a proper measurement system, and lack of experience or models for success (Sotirakou and Zeppou, 2006). However, very few papers recognize that cognitive limitation may be a prime factor that undermines utilization of PMS in developing countries. One problem is that the bulk of theoretical and empirical researches on PMS utilization in the public sector have focused on public organizations in developed countries which is more familiar with the concept of measurement and where measurement of performance is an accepted practice. Attention in the latter context has been shifted from answering the question of what is performance measurement to how it can be implemented and what are the conditions that are required for success (Berman and Wang, 2000). The significance of this factor is that cognitive limitation is not curbed to adoption but also to the actual use of measures which will depend on the capacity of intended users to understand and apply knowledge for performance improvement (Hawke, 2012). In line with Behn (2003), this study adds to the voice that key stakeholders would need to form a tentative theory about measurement and how it can be employed to foster improvement. It shows that there is a need for re-educating key stakeholders, particularly the leadership and management, about the importance and use of PM as an influential instrument for positive change to modernize and improve the public sector. The paper suggests that public managers need to see examples to understand the usefulness of measurement and need to be involved in the designing process.

PI Demand and Supply

Analyzing the sources and types of performance information demand and supply in the Yemeni public sector discloses two interrelated issues.

First, it reveals that performance disclosure in developing countries is mostly driven by external rather than internal political pressure. In line with Tillema, Mimba and Helden (2010), this study empirically suggest that much of the information produced in developing countries reflects the interest of more powerful groups such as funding bodies- mainly donors and financing agencies- while demand from other stakeholders seems to be ignored or have lower priority. It shows that not only are external stakeholders more interested in performance information than internal stakeholders but, equally important, that internal stakeholders- mainly management- are reluctant to reveal any unnecessary data. This supports the claim that public agencies are more inclined to use performance information for external reporting than for performance improvement (Taylor, 2011b).

Second, the findings show that there is insufficient demand for performance information and performance enhancement in the Yemeni public sector. The study supports the linear relationship between the demand for and production of performance information, and suggests that insufficient demand leads to limited supply. Both low level demand and supply in the public domain could be attributed to the weak accountability culture that seems to characterize the majority of public organizations in Yemen. While present legislations and administrative mandates may create important processes such as the development of annual reports, they do not seem to set out key principles for determining responsibility for performance. And, based on the analysis, accountability for results could not be determined based on annual reports, possibly due to the subjective nature of performance assessment, the lack of specific guidelines on how to prepare performance reports, and the liberty given to managers to show only their success in managing performance.

One possible conclusion from these two issues is that public organizations in developing countries aren't expected to adopt PMS by choice- as an internal decision- but rather by some pressure from outside. While many papers suggest that performance measures are adopted for their value in decision-making and that an internal policy requirement for change might have a stronger effect on adoption (Julnes and Holzer, 2001; Torres, Pina and Yetano, 2011), this paper alternatively suggests that the introduction of PMS in public organizations in developing countries is less likely to be the choice/decision of internal actors. It indicates that given the limited perception about the value of measurement, subsequent low demand for performance information, weak accountability culture, and low involvement of stakeholders, the chance of voluntary adoption of measurement is rather slim in the context of developing countries. This is not a major revelation if we consider the early stages of PM in the public domain. Many

authors note that measurement initiatives in many developed countries have been largely credited to external influence (Melkers and Willoughby, 1998; Bourdeaux and Chikoto, 2008; Cavalluzzo and Ittner, 2004; Degroff et al., 2010; Moynihan 2004; Ohemeng, 2010; Pollanen, 2005). Various legislations, administrative mandates and many other initiatives were based on the assumption that mandated reporting of results and performance indicators will lead to improvement in efficiency, effectiveness, and better accountability in public organizations. Consequently, as many governments in the world do legalize principles of performance measurement and management, the Yemeni government is advised to do the same if it is to institutionalize a PMS and reap the benefits of measurement for reforming the public sector.

Additionally, this paper highlights the need to identify champions for change in the public sector. It confirms the power distance between different stakeholders and concludes that any attempt to adopt PMS in the public sector should take into account different power positions by different stakeholders and study their influence. The paper empirically confirms the finding of Mimba, Helden and Tillema (2013) that funding bodies are the dominant stakeholders with the strongest power position to induce change in the least developed countries.

However, while this study shows that external pressure might be helpful for initiating change, it also illustrates that external pressure alone is not sufficient for success. Even if PMS was adopted to collect and distribute performance information, there is little indication that it will be used for decision making purposes in public organizations in developing countries. Findings show that even when managers would have to comply with external requirements for reporting performance they rarely depend on information from these reports for future planning or decisions. The limited use of performance data- regardless of its quality- indicates that PMS might not be adopted for its value but rather to gain legitimacy or financial advantages. Consistent with institutional theory, mandated organizational reforms in the public sector might end up being symbolic with little impact on internal operations (Cavalluzzo and Ittner, 2004; Greiling, 2006; Torres, Pina and Yetano, 2011; Ohemeng, 2011). To that end, this paper supports the conclusion of Sanger (2008) who maintains that legislations and external mandates can create the first step forward- a process for measurement- but the use and the value of the process is less certain.

Main Barriers

This study draws upon the experience of public organizations in Yemen to examine additional challenges that limit the implementation and use of PMS in developing countries. Five main factors have been identified that relate to the capabilities and culture of public organizations.

To start with, findings of this research show that lack of proper resources is a major challenge in developing countries. Because of limited finances, public organizations have been functioning under fiscal constraints and forced to do more with less. According to Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel (2006),

undeveloped economies and fiscal crises are common problems known to limit government programs in developing countries. However, while the literature indicates that adequate financial resources are essential to encourage efforts that aim at developing measurement (Van Dooren, 2005; Julnes and Holzer, 2001; Mimba, Helden and Tillema, 2013; Northcott and Taulapapa, 2012; Sole and Schiuma, 2010; Yang and Hsieh 2007; Lilian Chan, 2004; Rantanen et al., 2007), this study suggests that insufficient finance is not necessarily a barrier for adoption. Bearing in mind the lack of awareness and understanding of the value of measurement, most managers were not able to explain how budget constraints could prevent adoption. In the same way, the study did not find evidence that suggests a direct link between financial deficiency and actual use of measurement - possibly due to the limited experience of public organizations in Yemen in the field of measurement. Nevertheless, this research notes that lack of resources could affect implementation of measurement indirectly in two ways. First, as indicated by some interviewees, lack of finance could negatively influence training programs in public organizations. Our interviews revealed that because of limited budgets very few public organizations show deep interest in training their workers and managers are not bothered with designing training programs that fit their needs.

Consequently, the limited capacity of workers- which was cited as one of the main barriers- appears to be a chronic problem that is not expected to be resolved in the near future due to limited finance. The paper suggests that two levels of training could be adopted. First, government executives and workers would need to be trained in order to improve their cognition about the need and value of measurement. Second, training should cover the design of a proper system for measurement and selection of measurement. Training could also be used for defining goals and targets as well as communicating performance expectations.

The second way in which inadequate finance could influence utilization of measurement is through its impact on rewards. Managers were of the view that for PMS to work in public organizations there should be a strong incentive system that rewards and sanctions good and bad performance respectively. According to them, PMs could lose its meaning if it is not linked to the work of people. Their response accords with the view of Lee and Cho (2011) who suggest that setting the link between the two is at the heart of how measurement can lead to better organizational performance. The study shows that inadequate finance could threaten reward and incentive schemes in public organizations which mean a less motivated workforce that doesn't see the value of performance information.

Further, this study shows that ambiguity of performance objectives could be one of the main factors that impede the development and use of measurement in public organizations. Findings support some researches by revealing that objectives of public organizations in developing countries are mostly abstract and vague (Mimba, Helden and Tillema, 2007). And for the majority of public organizations there is no opportunity for benchmarking- no standards against which performance can be compared. The paper

concludes that public organizations with more tangible outputs and standardized services expected to be better able to assess their performance than others, and consequently are more likely utilize measurement. It supports the conclusions of Berman and Wang (2000) who maintain that organizations' ability to meet objectives is an important aspect for both the adoption and use of PMS. Yet, it is not clear as to why there have been limited attempts to clarify objectives in the public domain. Ambiguity could be intentional or unintentional and can be used to limit conflict or to avoid accountability and responsibility for bad results (Taylor, 2011a). Ambiguous goals could also lead to dysfunctional performance measurement because there will be multiple ways to report how goals are achieved (Goh, 2012). The study argues that understanding the reasons will help in forming an idea about the limited application of measurement in these organizations.

Furthermore, findings of this research support prior observations that fear of change and subsequent resistance to it may be a fundamental obstacle to the development and use of measures. It shows that managers and employees who are familiar with current traditional norms and practices may fear change or see no value in it and consequently be reluctant to adopt a new approach. The study backs other studies on the importance and the need to develop a supportive culture that is open for change and learning (Bevanda, Sinković and Currie, 2011; Goh, 2012; Julnes and Holzer, 2001; Taylor, 2007). It notes that collecting performance information provides an opportunity for decision makers to learn more about their activities and operations in order to improve performance and cut costs (Taylor, 2011b). The paper suggests that without a suitable organizational learning mechanism the value of measurement will continue to be limited in the public sector and resistance to change will be higher.

Several authors refer to corruption as an element that hinders reform initiatives and limits the potential for adopting measurement in the public sector, particularly in developing countries (Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006; Mimba, Helden and Tillema, 2007; Mimba, Helden and Tillema, 2013; Verheijen and Dobrolyubova, 2007; Bevanda, Sinković and Currie, 2011). Nevertheless, this research could not possibly establish whether there was evidence of extensive corruption where resources may be diverted for personal gains. What is more evident and relevant though is the absence of certain guidelines for reporting performance in addition to the lack of feedback on these reports from upper level government. The study shows that weak accountability culture might be more relevant than corruption. Many interviewees admit that there is hardly any action that has been taken to hold people accountable for their actions for quite a long time. This shows that reporting is a symbolic practice and reveals the weak accountability culture in the Yemeni public sector. The study concludes that internal and external stakeholders should work to enhance accountability by getting involved in determining the type of data that fit their needs. They should also define the quality of information they require and show their willingness to use that information for decision making and performance improvement.

Finally, although the majority of respondents did not refer to the lack of administrative leadership as a major barrier, this study however is of the view that it is a major challenge to the utilization of PMS in developing countries. For example, the inability of public officials to reward or sanction good and bad performance respectively, due to what seems to be their inability to allocate and manage financial and human resources for public programs, indicates that public officials have limited control over their environment. It is the conclusion of this research that managers who lack the authority to make decisions based on performance have little incentive to support measurement initiatives. Hence, if the influence of political leadership could be measured by the ability of political actors to mitigate and address technical problems and their ability to alter or shape the culture of their organizations as suggested by Moynihan, Pandey and Wright (2011), it should be expected that the influence of leadership in developing countries is rather limited because of certain restrictions on their ability to act. Further, this paper empirically supports the conclusions of Hawke (2012) who maintains that in the absence of external scrutiny of performance information, self-protection that is reinforced by the existing political culture will prevail at the expense of an organization's ability to meet their objectives and targets. The paper concludes that without proper consideration of these issues political actors not only lack the motive for developing and using measures for performance improvement but will also continue to resist any mandate for change.

Note about barriers

The analysis of barriers not only highlights the significance of adjusting systems of performance measurement and management to the specific cultures in non-western countries but more importantly to recognize the considerable differences between different types of public organizations. According to Julnes and Holzer (2001), one reason for the lack of influence of external mandates on the implementation of measurement is that such policies often disregard an organization's ability to implement change. Findings reveal that public organizations are different in many aspects, such as services, goals, functions, structure, source of finance, and management. For instance, notwithstanding insufficient resources in most organizations in the public sector, data indicate that public organizations do not face the same level of financial difficulties. The interviews revealed that some organizations are better financed than others, particularly those charging money for their services or receiving support from donors. Similarly, data suggest that not all aspects of performance in the public sector could be determined or measured in every organization. The issue is more evident for central agencies such as ministries that have problems in identifying their targets and outputs. Findings also show that some organizations are more open to change than others. While the majority of public organizations place little emphasis on learning from performance results and feedback, and current administrative mandates on reporting performance are not followed by formal rules and procedures to measure and report performance, there were some exceptions. The analysis indicates that, due to the conditionality of aid, it

seems that public organizations that engage with donors have stronger evaluative culture- a culture that encourages a deliberate quest for performance information in order to learn how to further enhance and improve performance.

Consequently, this study shows that there is a need to do more empirical work to understand how relevant these factors are to their context. It demonstrates that examining the public sector as one homogenous group is an unwarranted oversimplification made by many in the research field. This should indicate that a one-size-fits-all policy should be avoided by policy makers who should be mindful of possible dissimilarities between public entities. The paper suggests that setting up a measurement policy should take into consideration organizational characteristics that foster or impede the success of such a policy. This means that the direction of future endeavors to utilize measurement in public organizations should work to determine the level of readiness for change as suggested by Julnes and Holzer (2001). The study indicates that an organization's readiness could be determined by learning about the following: level of knowledge about the use and usefulness of measurement amongst key stakeholders; sources and types of demand and supply of performance information; availability of adequate financial and human resources; organizational culture and openness to change; and the level of support for measurement.

Last but not least, one more problem with past research is that it did not investigate the relationship that might exist between important factors and problems. Based on findings of this research, factors do interrelate with each other and addressing basic challenges might mitigate smaller ones. As discussed earlier, availability of financial resources could have a positive or negative impact on training and incentive schemes. The paper concludes that investigating the link between problems and the type of organizations could be helpful to explain not only how certain challenges happen in reality but, equally important, why they happen in the first place. Further, given the differences in the implementation of measurement in various organizations, investigating the link between problems and the type of organizations could be helpful to determine generic problems as suggested by De Waal and Counet (2009). It is to that end that this study calls upon future researches to be mindful of the interrelationships that might exist between barriers and the significance of their interaction over the utilization of measurement.

Conclusions

This paper examined the patterns of PM usage in the Yemeni public sector by relying on the views and experience of various public managers in Yemen. The aim of this research is to offer an explanation for the slow implementation and disappointing effects of PMS in the public domain within the context of developing countries. While empirical investigation is focused on the Yemeni context, the overall analysis could have a general validity. The study highlights fundamental issues related to the development and use of PMS in public organizations in non-western countries. It focuses on the important conditions needed to ensure the effective utilization of measurement in the public sector.

Findings indicate that the Yemeni public sector features a traditional administration that is away from a result-based management, bound to roles and process, and less flexible to change. Consecutive reforms in the public sector did not seem to emphasize the need to establish a performance-based culture and accountability for results.

The Yemeni experience offers several valuable lessons. It suggests that there is a cognitive limitation to the adoption of PMS in developing countries where main stakeholders are less familiar with the concept of measurement and do not see its value. The study suggests that given the limited perception about the value of measurement, weak accountability culture, subsequent low demand for performance information, and low involvement of stakeholders in developing countries, anticipating an optional adoption of PMS in which internal stakeholders make the decisions for change is a bit overoptimistic. It argues that integrating performance measures into the management process is less likely to happen without external intervention. The study argues that institutionalizing principles of performance measurement and management through legislations might be helpful for initiating change in the public sector, particularly in developing countries. The paper also proposes that mandating performance measurement can enhance knowledge about the use and usefulness of measurement by means of practice- learning by doing. It shows that legal mandates to use measures in developing countries could also increase demand for performance data and enhance accountability in the public domain.

However, some evidence in this research shows that institutionalizing reform will likely end up being symbolic for legitimacy purposes only and not for performance improvement. The fact that performance information is mostly produced in response to external rather than internal pressure, and that managers rarely depend on performance data to make decisions, support this conclusion and show that while initial adoption could be achieved by outside pressure, proliferation of measurement- further adoption and actual use- is a different story. The paper calls for a balanced and value focused approach in order to reap the benefits of measurement. A balanced approach suggests

that the instrumental benefits of measurement should be as important as the symbolic benefits. This means that, in line with external interest in performance information, public organizations should ensure internal buy-in (commitment) to make measurement work. Internal and external stakeholders should recognize the value of measurement and should not feel intimidated by change.

Additionally, and perhaps most importantly, the study also highlights the importance of taking organizations' differences into account when it comes to the utilization of PMS in the public sector. Treating public organizations as a homogeneous group that has the same functions, characteristics, and conditions seems to be a common oversight in past researches. The paper suggests that utilization of PMS is less likely to be neutral and is conditioned by environmental context of each organization. It concludes that the influence of internal and external requirements for measurement in the public sector could be limited by certain elements that are related to organizational capacity and the dominant culture in developing countries. The study empirically demonstrates some of the elements that explain differences between public organizations. It claims that these elements can be used to determine the level of readiness of a particular organization for measurement. For example, this paper argues that public organizations that have adequate finance, clear objectives and services, trained and qualified workforces, adequate incentive systems, a supportive and open culture for learning and improvement, and most importantly administrative leadership that is motivated and fully committed for change are more likely to utilize measurement successfully than others.

However, while findings of this paper may indicate that utilizing PMS in line agencies will be far easier than in central agencies, it also shows that changing predetermined conditions and the current culture to fit the need for measurement in the latter is not an easy mission. It requires the development of institutional capacity that includes current working practices and processes and re-building the skills of management and workers. Constant political support and considerable resources would be needed to foster change and transformation. This means a complete shift in organizational culture to change current attitudes and working habits towards result-based performance.

Limitations and Further research

This study reveals that further research is needed to understand the conditions that are needed to make measurement work in public organizations, particularly in the developing world. As explained in the literature review, up until now little was known about the organizational and contextual factors that determine the success or failure of the utilization of PMS in the public domain. Assuming that PMS can enhance efficiency and effectiveness and can be readily transferred to public organizations in developing countries is a bit overoptimistic. Future empirical studies and concurrent theoretical testing would provide useful guidance to harness the potentials of measurement to reform the public sector and increase its efficiency and effectiveness.

Limitations affecting the generalization of the findings must be noted. While some issues are relatively generic- applicable to the majority of public institutions- some issues are more specific than others, which make generalization difficult. Accordingly, this study highlights the importance of taking organizations' differences into account when it comes to the utilization of PMS in the public sector. It implies that utilization of PMS is less likely to be neutral and is conditioned by the environmental context of each organization. Findings support the idea that government organizations are different in many aspects such as goals, tasks, structure, funding, capacity, political environment, and types of purchasers of public services. Hence, whether findings are mirrored across the Yemeni public organizations will only be determined by further researches that include different type of organizations in the public sector. Further, while a larger and random sample that include participants from different kinds of public organizations may seem ideal for generalization purposes it is difficult to achieve in reality, particularly under current conditions in the country. Furthermore, findings are based on the views of public managers and it is not known to what extent their views reflect reality. While interview questions were set to solicit managers' understanding of performance measurement and management, it is possible that they may have different interpretation of the concept than that in the management literature. In the same way, relying on the opinions of officials alone may not provide sufficient information about the context and the factors that distinguish between public organizations in relation to the adoption and use of measures.

This research may be advanced in multiple ways. First, it could be advanced methodologically by adopting a comparative case study method on a number of different public organizations whether nationally or internationally. Comparative case studies are useful to discover how two or more cases compare in specific outcomes, conditions and criteria. It should allow further investigation and scrutiny of different processes and behaviors in different socioeconomic and cultural settings in order to identify common elements from those specific to their context (Bianchi and Rivenbark,

2012). Further, since the sole reliance on the input and narrative of officials may not provide enough information about the context and the elements that distinguish between public organizations in regards to the utilization of measurement, data collection could also be enhanced by examining available documents of each case. Future research could also examine behaviors, motivations, and relationships in different settings, and should put more emphasis on the relationship between problems in order to come closer to a possibility to generalize findings.

Second, the paper highlights the need for improving theorization about why and how measurement policy can be adopted and used outside the private sector. Findings of this research should have worked to found the theoretical position that is needed to direct further data collection and analysis in the case study approach (Gray, 2004). The relation between demand and supply of performance data and the subsequent impact on the actual use of measures as suggested by this study could be useful in that direction. For instance, future research could further our understanding about the ways that public organizations respond to pressure from their institutional environment and adopt structures and practices based on expectations and social value (Torres, Pina and Yetano, 2011).

Finally, given the limited research in this area, this study is with the view that similar researches should be repeated periodically in order to update the literature with the list of issues and challenges that influence the utilization of PMS in public organizations especially in developing countries. Conclusions should be studied, tested and enhanced further by future empirical research based on data collected from other countries.

APPENDIX 1 – INTERVIEW SCHEDULE

General Information

Name	Organizational Level	Level of Education	Experience	Type of public organization	No. employees
------	----------------------	--------------------	------------	-----------------------------	---------------

Managers' Perception

The first part should focus on eliciting how public managers perceive measurement of performance and whether they are familiar with the concept. It should determine whether these managers use any sort of measurement or indicators of performance for any management purposes or functions. It also explores their views about the usefulness of measurement in general and the role of measurement in boosting organizational performance.

How do you manage performance? Or, how do you know whether your organization is performing well or not?	كيف يتم ادارة الانجاز واداء العمل؟ او كيف تعرف ان المؤسسة تحقق اهدافها؟
(Probe for examples. Ask about types of outputs, products or services? What is the desired outcome? Who are the customers? Etc.)	

Do you have a systematic process for measurement? Do you use any type of measures or indicators?	هل تستخدم اي مقاييس او مؤشرات؟
Can you give me examples?	الرجاء اعطاء امثلة على ذلك؟

Can you link your performance to the overall objectives and strategy, and how?	هل باستطاعتكم ربط العمل الذي تقومون به مع الاستراتيجيات والاهداف العامة للمنظمة، وكيف تقومون بذلك؟
--	--

Do you think measuring performance is good for you? Why?	هل تعتقد ان قياس الاداء مفيد بالنسبة لكم؟ لماذا؟
How would you describe the use and purposes of performance measurement? How could it help you?	
كيف تصف فوائد واستخدامات مقاييس العمل؟ كيف لها مساعدتك في عملك؟	

Types and Source of PI Demand and Supply

This part is concerned with identifying the sources and types of performance information demand in the public sector. It should work to reveal the relationship between interviewee's organizations on the one hand and internal and external stakeholders on the other in terms of reporting performance.

Whom do you report your performance to, and why?	الى من ترفعون تقارير انجاز العمل ولماذا؟
Is it optional or mandatory to report your performance? Is there any legal or administrative mandate for reporting performance?	
وفيما اذا كان رفع مثل هذه التقارير اجباريا ام اختياريا؟ وهل هناك متطلبات قانونية او ادارية لرفع تقارير الاداء؟	

Can you explain how do you receive request/demand for PI, and how do you report performance?	كيف تتلقون استفسارات بشأن الاداء؟ وكيف تجيبون عليها؟
--	--

Can you identify users of PI inside and outside your organization?	هل بالامكان تحديد المستخدمين داخل وخارج المؤسسة للمعلومات المتعلقة باداء العمل؟
--	---

Can you identify and explain any other internal or external requirements for reporting organizational performance?	برجاء تحديد وشرح اي متطلبات اضافية داخلية او خارجية لرفع تقارير الانجاز والعمل
--	--

Main Barriers

The final theme was focused on understanding possible barriers and limitations to the utilization of measurement. It encourages respondents to share their experience in relation to the development and uses of measurement or to give their opinion about possible hurdles in the process of adoption.

What type of difficulties (if any) do you encounter when it comes to meeting your objectives?	مانوع الصعوبات (ان وجدت) التي تواجهونها لتحقيق اهدافكم؟
According to you, can you measure performance at all times? What factors, and to what extent, impede the development, use, and reporting of measures?	برايك هل بالامكان تقييس العمل في كل الاحوال؟ ماهي العوامل التي تعيق تطوير واستخدام مقاييس العمل والى اي مدى؟
In your opinion, what are the requirements for proper PMS?	برايك, ماهي المتطلبات اللازمة لنظام ملائم لقياس الانجاز والاداء للعمل؟
Do you think other countries have been more successful in implementing and using PMS in public organizations? Why?	هل تعتقد بان دولا اخرى في الخارج افضل في تطبيق واستخدام نظم قياس مستوى الانجاز واداء العمل؟ لماذا؟
According to you, what could motivate management to adopt and use performance measurement?	برايك, ما الذي قد يدفع او يشجع الادارة في مؤسستك على تبني واستخدام مقاييس الانجاز؟

**CHAPTER 6: CONCEPTUAL FOUNDATION FOR A THEORY OF PMS
UTILIZATION IN THE PUBLIC SECTOR**

Overview

This part of the study is intended to present the theoretical contribution of this research in forms of a new conceptual model for the utilization of PMS in the public sector. The chapter is organized as follows.

The first part describes the theoretical foundations of the proposed model. This entails a general overview about the rational and background of model building- why and how it was developed.

The second part of this chapter provides a general description and graphical representation highlighting the main elements that distinguish this model from others.

The third and final part provides an extensive description of the main building blocks of the conceptual model. It begins with a detailed description about the significance of looking at two different levels of PMS utilization: adoption and implementation, followed by a discussion about the need to look into the two levels of utilization in terms of processes within stages. The assumptions that lead to the inclusion of these processes and the definition of each are explained in that section. The section that follows will provide a detailed description about the main factors that influence utilization and how they affect each process. And, the final section will introduce a general discussion about possible interdependence and interrelation between processes, between factors, and between processes and factors. The overall discussion will be conducted in a way that show what parts have been influenced by the literature and what have been shaped by the empirical findings of this study.

Theoretical Background

The conceptual model builds and extends on the framework proposed by Julnes and Holzer (2001) which later was applied and tested by Carlucci, Schiuma and Sole (2015). Julnes and Holzer built their original framework by focusing on the utilization problem and previous findings that suggest a need to understand: why some public organizations adopt PMS while others don't; why some organizations have been only selective in the development of measures; and most importantly why the development of measures doesn't guarantee the actual use of these measures. Accordingly, Julnes and Holzer suggested that utilization is a change process that involves certain stages that need to be recognized and supported. Reflecting on previous studies, they viewed utilization of PM as consisting of two main stages- adoption and implementation. Their study employed the results of a national survey of state and local government entities to investigate the

factors that affect the utilization of performance measurement. Findings were also used to develop a theoretical framework that explains how those factors influence utilization Figure (17). Authors adopted a rational/technocratic and political/cultural view to provide a thorough panel of factors that influence utilization of PM in the public sector. The following conclusion was reached:

“Policy adaption is driven more heavily by factors from rational and technocratic theory, whereas actual implementation is influenced by factors addressed by political and cultural considerations” (P 693).

Carlucci, Schiuma and Sole (2015) on their part worked to address the noted gap between the development of performance measures and their effective use which, in their view, is one of the major causes of failure of PMS. They built on the original framework of Julnes and Holzer to introduce a graphical conceptual model for the purpose of testing several hypotheses related to the relevance of factors on each stage of utilization. Using the labels of Julnes and Holzer, their new model proposes that factors affecting adoption and implementation of PMS can be labelled as rational, political, and cultural factors. The model was tested by using structural equation modelling on a survey data from an Italian region and district to reflect the cause and effect relationships between seven key sub-factors- pertaining to the abovementioned main factors- and the two stages of utilization- adoption and implementation Figure (18).

Accordingly, the relevance of all elements whether in the original framework of Julnes and Holzer or the modified model of Carlucci, Schiuma and Sole has been examined against the findings from the systematic review and the empirical work of this study leading to key changes that will be discussed extensively in the course of this chapter. The discussion will explain which parts were affected by the literature and which have been influenced by findings of the empirical work.

	Adoption	Implementation
Rational/technocratic	External requirement Internal requirements Resources Goal orientation Information	Resources Information
Political/cultural	Internal interest groups	External interest groups Percent unionized Risk taking Attitudes

Figure 17 Adapted from Julnes and Holzer Framework 2001

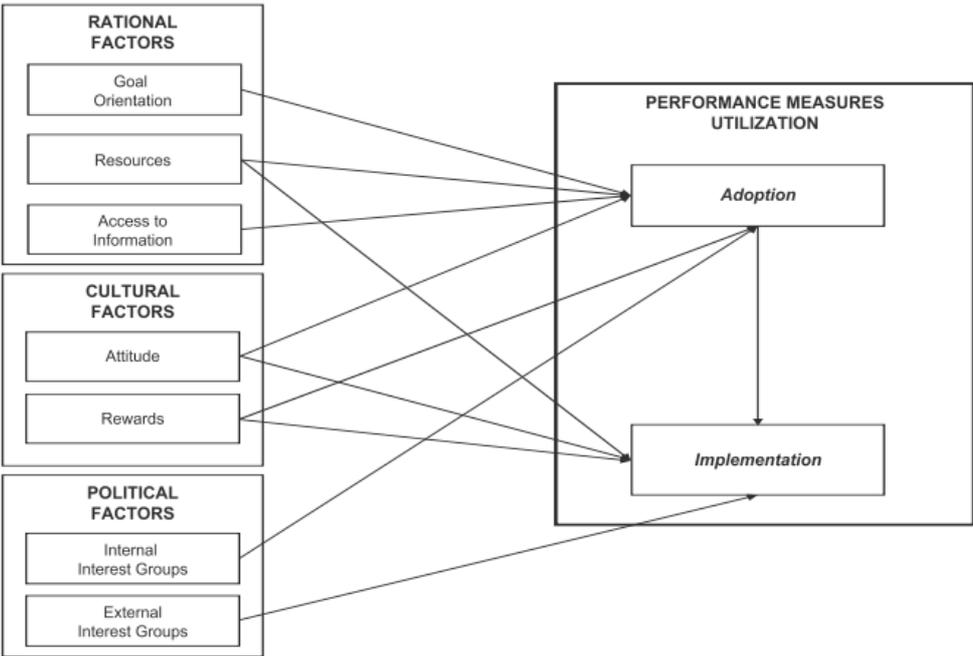


Figure 18 Adapted From Carlucci, Schiuma and Sole’s 2015 Model

General Overview of the Conceptual Model

Referring to the abovementioned work of Carlucci, Schiuma and Sole (2015) and Julnes and Holzer (2001), the development of a theoretical model that can be used to summarize the main elements of my reasoning regarding utilization of PMS in the public sector was also guided by the work of (Whetten, 1989) who indicates that a complete theory/model must consider the following building blocks:

- WHAT factors should be considered? based on judging comprehensiveness versus parsimony criteria;
- WHY they should be considered? based on identifying main assumptions;
- HOW they are related? By showing causality and delineating patterns.

While Whetten argues that the ‘What’ and ‘How’ questions constitute the domain or subject of the theory, he also notes that:

“The more complex the set of relationships under consideration, the more useful it is to graphically depict them”

(Whetten, 1989, p.491).

The result from synthesising the theoretical and empirical findings of this study with the work of (Carlucci, Schiuma and Sole, 2015; Julnes and Holzer, 2001) and the guidelines and principles of Whetten (1989) is the model illustrated in Figure (19). The model shows three different processes (demand, supply and use) at two levels of utilization (adoption and implementation). It also shows main factors affecting PMS utilization in the public sector. The model illustrates how these factors relate to one another in shaping the adoption and implementation of PMS and possible inter-play between processes and factors and the level of significance at each level of utilization.

Circle shapes on the right side represent processes of utilization at the two levels adoption and implementation. The adoption level is represented by two main processes demand and supply, while the implementation level is represented by one process that is concerned with the actual use of measurement information for decisions’ purposes. The interaction between these processes is represented by dark black-coloured arrows. These arrows are designed to show the constant interaction between processes at each level of utilization. Two-way arrows signify two direction of influence in which one process leads to the other and vice versa.

The three main factors derived from this research are delimited by large rectangular boxes on the left side. Smaller rectangular boxes represent the most relevant sub-factors identified by this research for each category of these factors. The interconnection between main factors is represented by dark solid arrows each with distinct colour to

exhibit the influence of each factor independently. For instance, external political influence is represented by dark orange-coloured arrows to distinguish it from the influence of internal politics that is presented by dark blue-coloured arrows. The model uses dotted arrows to show the inter-dependence between sub-factors pertaining to organizational capacity and organizational culture. These arrows show that this is an interdependent relationship in which main organizational culture attributes are dependent on organizational capacity elements. Training is a mediator sub-factor that is represented by an oval shape which is affected by and dependent on other sub-factors.

Finally, a cloud shape is used to represent the perceived value of measurement to external and internal stakeholders. It represents the principle motivation- instrumental or symbolic- for the development and use of PM. Loops of dotted arrows are used to reflect feedback and how perceived value could change based on experience and reflection from the supply and use processes and based on training outcomes. One-way arrow signifies one direction of influence and two-way arrows emphasize reciprocal influence. Reflecting on the case findings, the model also suggests that a cognitive barrier may prevent internal or external stakeholders from seeing the value of measurement which in turn undermines the demand for measurement information. The cognitive barrier is represented by vertical, thick and dotted line.

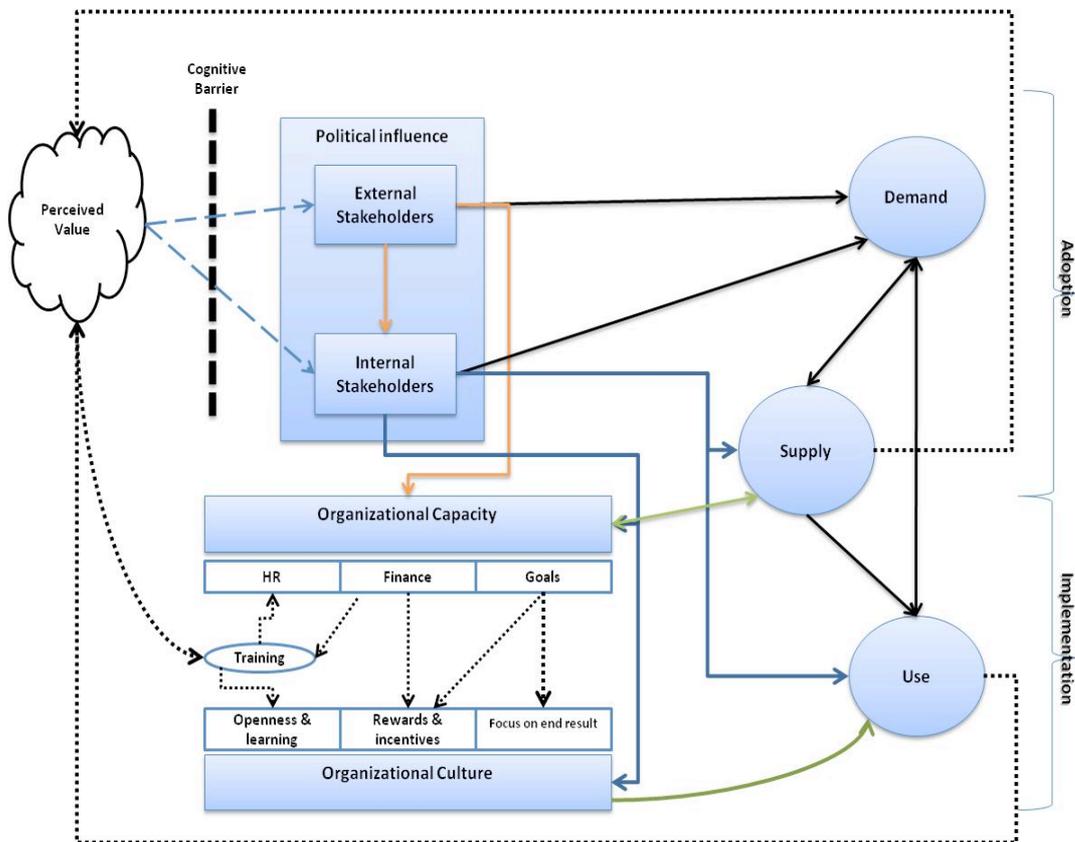


Figure 19 the Conceptual Model

The following table presents a summary of the main elements that distinguish my conceptual model from those of Julnes and Holzer (2001) and Carlucci, Schiuma and Sole (2015).

Table 18 Theoretical and Empirical Findings

Previous Framework/Model	My Conceptual Model
Utilization of PMS is a change process consisting of two stages: adoption and implementation	Utilization of PMS is viewed in terms of different processes at each stage. Adoption is viewed in terms of two processes- demand and supply; Implementation is viewed in terms of one process- the actual use of PI to make decisions.
Utilization of PMS is affected by political, technocratic/rational, and cultural factors. Technocratic factors: goal orientation, resources, and access to information. Cultural factors: attitudes and rewards.	Utilization of PMS is affected by political factors, technical capacity, and cultural considerations. Capacity issues: HR, Finance, and clarity of goals. Cultural factors: openness to change and learning, rewards, focus on end results.
Do not consider how the perceived value of measurement affects utilization	Different processes of utilization are affected by the perceived value of measurement.
PMS is common knowledge	Recognizes the lack of understanding of the use and usefulness of measurement as a possible barrier particularly in developing countries.
Solely focused on how factors affects utilization with no proper consideration of the interrelation between factors or between processes	Provides a description about possible interrelation between processes, between factors, and between factors and processes.
Based on deductive approach with quantitative strategy and based on surveys for data collection	Based on inductive approach using qualitative strategy and live interviews for data collection

Elements of the Conceptual Model

This section is intended to present the main elements of the new conceptual model illustrated in Figure 3 along with rationale for major adjustments on the original framework and model of Julnes and Holzer (2001) and Carlucci, Schiuma and Sole (2015) respectively. The model can be described in terms of three main building blocks: levels and processes of PMS utilization, displayed in the right side of the graph; factors and sub-factors that affect PMS utilization, shown in the left side of the model; and interrelations between different components of the proposed model. The relevance of other elements such as the perceived value and cognitive limitation will also be discussed in the course of this section.

The Utilization Issue

The first building block of my conceptual model starts by learning what PMS utilization really means. The paper answers the question by first recognizing the value from identifying different stages of utilization and second by presenting the rationale for using processes in addition to stages.

Two levels of Utilization

As stated earlier, this model reflects on the work of Julnes and Holzer (2001), Van Dooren (2005), and Carlucci, Schiuma and Sole (2015) to define utilization in terms of two main stages: adoption and implementation. Adoption is about having performance measurement or the extent to which an organization uses measures to determine its own performance. It represents the first step of ‘developing a capacity to act’ or ‘generating knowledge’ which, for example, involves the development of measures of inputs, outputs, efficiency, and outcomes. Implementation on the other hand, refers to the way an organization uses measures for its own benefits. It is about ‘doing performance measurement’, or ‘knowledge converted into action’ which is simply the actual use of performance measures to take certain actions or decisions for various purposes such as, to name a few, planning, allocation of resources, reporting, and evaluation.

Several lessons can be learnt for both practitioners and academics alike from recognizing these two stages of utilization. Firstly, as noted in the Systematic Review, adoption and implementation is not the same thing and consequently determinants of one are not necessarily determinants of the other (Yang and Hsieh, 2007). Scholars need to form a better understanding about this distinction in order to avoid the problem of focusing on one stage and ignoring the other. Secondly, it becomes clear now that the mere development of measures does not necessarily guarantee the actual use of these measures which is why there is a need to isolate and determine the factors that influence each stage. Finally, identifying factors in relation to the two distinct, but interrelated, stages should be helpful for an effective utilization of PMS in the public domain. It provides valuable information about several important aspects such as: the level and sources of support for performance measurement and management, the organizational

readiness for adoption or implementation, and the usefulness of measurement information to improve performance.

Another implication concerns the interest of internal and external stakeholders involved during each stage. Different stakeholders have different interests and the literature informs us that the interests of both external and internal stakeholders are not always equal. For example, while adoption may be driven by external institutional pressure, implementation might be more influenced by internal stakeholders based on their perception of the value of measurement. Accordingly, it might be logical to assume that mobilizing political interests and gaining larger support is normal during adoption but as change takes place and programs get implemented the momentum of support and interests starts to dwindle. In short, the proposed model indicates that recognizing adoption and implementation as two distinct stages is useful to identify and isolate the effects of factors for each stage which is helpful to address the challenges for an effective utilization.

Utilization Processes

Building on my **synthesis** in the Systematic Review, this model takes a step further by explaining adoption and implementation of performance measurement in terms of three main processes: demand, supply, and use. This is not identical to the work of Julnes and Holzer (2001) and Carlucci, Schiuma and Sole (2015) who focused on the two stages of utilization only and not on the processes at each stage. While adoption is explained in terms of the first two processes, demand and supply, implementation is concerned with the use process. This arrangement makes a clear distinction between real interest in measurement, the development of measures, and actual use of measurement information.

The decision to use processes to study and explain the utilization of PMS is based on a thorough reflection on the literature and results from the empirical project. For example, Van Dooren (2005), Mimba, Helden and Tillema (2007), and Tillema, Mimba and Helden (2010) used the supply and consumption analogy to distinguish between the production process of information and the use of that information. The supply side represent the production process of PM and the demand/consumption side represents the actual use of performance information for internal management and control. Authors explain that the supply and demand approach to PM is embedded in public policy research and that performance information is viewed as one factor-amongst many- in decision making. However, defining demand in terms of the actual or desired use of performance information creates confusion because it is likely to become confused with other meanings of demand. For instance, many papers indicate that adopting PMS in public organizations is mostly driven by internal or external demand/pressure for performance information (Moynihan, 2004; Moynihan, Pandey and Wright, 2011; Taylor 2011a; Tillema, Mimba and Helden, 2010; Torres, Pina and Yetano, 2011; Mimba, Helden and Tillema, 2013; Sotirakou and Zeppou, 2006; Sanger, 2008; Yang

and Hsieh, 2007; Hawke, 2012; Brusca, Manes Rossi and Aversano, 2015). Consequently, the definition of ‘demand’ in those papers is not identical to that asserted by (Van Dooren, 2005; Mimba, Helden and Tillema, 2013). Further, the empirical work shows that one key to non-adoption of PMS in the Yemeni public sector is the lack of demand from the part of internal and external stakeholders. Suddaby (2010) points to that problem by explaining that:

‘When researchers use different terms for similar phenomena, it produces confusion “confounding effects” that impede the ability of members of a research community to communicate with each other or to accumulate knowledge’

(Suddaby, 2010, p.352).

Consequently, in order to avoid confusion and recognize the significance of political influence, this conceptual model uses **demand as a standalone process** to explain the level of involvement of stakeholders in PMS utilization in the public sector.

Since only few research projects have been focused on understanding how political agents perceive and use measurement (Ho, 2005), the proposed model suggests that the demand process can be used as an efficient means to monitor changes in the perception of value. In other words, the role of political actors and their perception of the tool can be detected by looking into their level of pressure for adopting and implementing measurement. When political actors are not interested in measurement they lose enthusiasm and it’s likely that their demand for performance data will be low as well. Therefore, demand shows that adopting PM is not distinct from the will of political actors inside or outside the organizations and stresses the role of power and politics in the overall process of utilization. It could be considered as a reflection of public officials’ concern about accountability, responsiveness, and quality of service. Additionally, demand for performance data determines the motivation behind measurement which in turn determines what types of measures need to be used. It can be used to establish the usefulness of measurement in terms of its ability to meet expectations, exceed expectations, or not meeting expectations at all.

The supply refers to the process of measurement or the actual development of performance measures. It focuses on determining whether an organization has been able to meet pre-established objectives using measurable standards for performance (Gao, 2009). As outlined by Fryer, Antony and Ogden (2009), four aspects could be noted for supply/performance measurement: deciding what to measure; how to measure it; data interpretation; sharing and communicating results. These are managerial choices that have significant implications on the utilization of measurement. The biggest challenge during this stage is related to the first two questions: deciding what to measure and how to measure it. It underscores the importance of the design of the measurement system and possible requirements for measurement. A framework of measurement indicates that PMs are utilized as means of describing organization’s strategy and objectives (Radnor and Lovell, 2003). Previous researches have shown that different organizations

would experience various systems for measuring performance with different models of measurement (Torres, Pina and Yetano, 2011).

Implementation is explained by one process only, **the actual use** of the information gained through measurement. Actual use is about what organizations decide to do with the knowledge it has acquired about its performance. It is characterised by the ability of stakeholders to make an informed decision(s) based on performance data. Ideally, accounting for current performance is expected to be followed by an action or decision for improvement. While adoption can be measured by the simple collection and presentation of performance information, the implementation phase is determined by the impact created by measurement information. The level of use of performance information can be explained in terms of the level of integration of that information in the actions of the organization. One way to measure that is by looking into the frequency of internal discussions of results and the attendance of management of these discussions (Van Dooren, 2005). Another way is by looking into the frequency of use of output and outcome measures to take certain decisions such as changing certain processes or allocating financial and human resources (Julnes and Holzer, 2001). ‘Actual use’ and ‘implementation’ refer to the same meaning and may be used interchangeably during the course of this paper.

Factors Affecting Utilization of PMS

Following an extensive review of performance measurement and management literature and in light of empirical findings, the proposed model identifies the following main factors that affect the utilization of PMS.

Political factors

Utilization of PMS in all organizations, public or private, would have to consider the influence of power and politics. The analysis of both the literature and the case findings suggest that stakeholders’ support and involvement is central during all stages of utilization. Both stages, adoption and implementation, could be considered as political choices made by individuals or groups that have the power to make decisions. It is related to the role of stakeholders that exercise some pressure and control over organizational resources and operations. Hence, understanding the political context of public organizations is necessary for successful utilization of measurement (Julnes and Holzer, 2001). The model consequently proposes the need to consider the influence of different political groups inside and outside public organizations. It makes a distinction between two types of political involvement external and internal. The significance of making a distinction between internal and external stakeholders is related to understanding their influence on each level of utilization. Such distinction recognizes that the public sector is characterised by shared responsibilities and co-production (De Bruijn, 2002). Multiplicity of stakeholders in the public sector signifies multiple motives, involvement and expectations. This adds certain complexities to adoption and implementation of PMS in that context.

External interest groups

External interest groups refers to all stakeholders outside public organizations who although not an internal member of the organization but have some influence and interest in its resources, operations and services. It encompasses all individuals and national or even foreign entities that exert some pressure on public organizations. Legislators, elected officials, citizens, upper-level government, labour unions, funding agencies and donors are all part of that category. Many papers noted that continuous support and involvement of external interest groups is important for successful utilization of performance measurement and management (Nath and Sharma, 2014; Degroff et al. 2010; Šević, 2005; Ohemeng, 2011; Mimba, Helden and Tillema, 2013; Verheijen and Dobrolyubova, 2007; Verbeeten, 2008). In fact, building on the experience of many public organizations around the world, many measurement initiatives were mostly attributed to external pressure (Ohemeng, 2009; Moynihan, Pandey and Wright, 2011; Mimba, Helden and Tillema, 2007; Cavalluzzo and Ittner, 2004; Hawke, 2012; Torres, Pina and Yetano, 2011).

Internal interest groups

Internal interest groups refer to internal members of the organization such as upper-level management, middle-level management, and workers. The influence and importance of internal actors is evident throughout the utilization process. This group of stakeholders play a key role in initiating change, managing the change process and creating the needed infrastructure for successful intervention. The proposed model shows a greater role for internal actors in all three processes of utilization. It demonstrates the role of internal stakeholders in initiating demand, designing and selecting measures, and using measurement information to alter change. Internal stakeholders are also better suited to determine the impact of the utilization process and reasons for success or failure. Hence, as noted by many, building a high level of commitment among upper-level management and gaining the support of workers is essential for success (Berman and Wang, 2000; Cavalluzzo and Ittner, 2004; Lilian Chan, 2004; Folz, Abdelrazek and Chung, 2009; Gao, 2015; Goh, 2012; Jackson, 2005; Liu, 2009; Moynihan, Pandey and Wright, 2011; Radnor and Lovell, 2003; Sanger, 2008; Yang and Hsieh, 2007; Brusca, Manes Rossi and Aversano, 2015).

The following section provides additional description about how the model views the impact of internal and external politics on each process and the relevance of the perceived value factor.

Political Influence and Perceived Value

To better understand how internal or external interest groups influence utilization, the model suggests a need to consider stakeholders' perception of value at each process.

It proposes that the demand process is a reflection of internal or external political interest in measurement. The final outcome of the demand process, according to this model, is not the development of measures but rather a decision to adopt measurement. Such decision depends on how political actors perceive the value of measurement. It is based on the assumption that the advantages of the new system outweigh the cost. The case study findings support the systematic review in indicating that when internal and external political groups realize the potential benefits of measurement they will be more motivated to adopt it. Accordingly, the proposed model suggests that a positive perception means that stakeholders will likely increase demand and maintain their support for measurement initiatives. Conversely, a negative perception means little interest in PM from the part of stakeholders which will decrease demand for performance information.

Based on the empirical work the conceptual model also suggests that lack of awareness and understanding of the benefits of PM represents a cognitive barrier that undermines or even blocks the demand process. While the PM literature seems to have fallen short in understanding how cognitive barrier affect adoption (Zeppou and Sotirakou, 2003), the case findings suggest that cognitive limitation may be a prime factor for the non-demand of measurement in governments within the context of developing countries. The model consequently reflects this and suggests a need to first remove the cognitive barrier- the vertical, thick and dotted line - in order to enhance the positive perception of measurement and boost demand for performance data.

The supply process is also influenced by the wishes and expectations of internal and external interest groups. Deciding what to measure in the public sector is dependent on the demand of stakeholders. However, the multiplicity of stakeholders in the public sector means that the value of measurement might not be the same for all. One possible implication for that is provided by Moynihan (2004) who explains that while internal stakeholders may be interested in the internal use of measurement to create instrumental value for their agency, external stakeholders maybe more concerned with the symbolic benefits that are based on the ability to communicate organizational rationality and accountability. Meeting the demand of all stakeholders might require different types of performance measures resulting in producing conflicting measures (Rantanen et al.

2007). Consequently, deciding what measures to use to make decisions would be difficult because of conflicting objectives of various stakeholders. Some researches even suggest that meeting the demands of multiple stakeholders will lead to a position where the measurement system is closer to form-filling than to provide useful information (Fryer, Antony and Ogden, 2009). In effect, this model recognizes that the first challenge during supply would be how to determine the purpose of measurement and how to establish a trade-off between competing objectives. Based on the case findings and some papers (Tillema, Mimba and Helden, 2010; Mimba, Helden and Tillema, 2007), it is logical to assume that the supply will likely reflect the interest of more powerful groups in public organizations and this is represented in my model.

Further, the model demonstrates that internal and external stakeholders influence supply indirectly through their control on resources. Political actors may have the power to mandate change but they have to consider organizational readiness for change in terms of capacity that is needed for the development of measures (the interrelation between political groups and organizational capacity and how that affects the supply process will be explained further in a latter section).

The model also demonstrates that internal political groups have more direct role and influence on the supply and use processes than that of external stakeholders. This can be noticed from internal actors' involvement in the development of measures and decision making based on performance data. Based on empirical findings and the literature, the responsibility for developing and gathering information falls on organizational members and actual use occurs when internal interest groups, mainly managers, understands the value of performance information and when non-management employees are open to the use of measurement for making decisions. Conversely, the model suggests that except for their direct influence on demand and their influence on resources, the involvement of external interest groups during supply is partial. The data from the research show that external stakeholders had little knowledge and interest in PM initiatives and that their commitment to the development of measures is limited. Similarly, given that the development of PM requires considerable time and that many external political groups are motivated by short-term concerns (Cavalluzzo and Ittner, 2004), it is expected that external political influence on implementation would also be low. While this is against the correlations predicted by Julnes and Holzer (2001) and Carlucci, Schiuma and Sole (2015), the case findings and other studies support that conclusion of Folz, Abdelrazek and Chung (2009) and Taylor (2011a). One possible justification for variation is the data collection method. Both studies of Carlucci, Schiuma and Sole (2015) and Julnes and Holzer (2001) depended on survey researches to solicit the views of officials and executives in several public organizations. One noted limitation of that method is the possibility for respondents to misinterpret questions and associate different meaning to them from that of the researchers.

The dynamic of political influence

This model can be employed to address one of the limitations of past research regarding the dynamism of political support. Building on the remarks of Bourne (2005), this conceptual model shows that political support and commitment is not constant but rather dynamic and continues to change over time. It is linked to the perceived value of measurement and the way the measurement system can serve the interest of stakeholders. Accordingly, the model reflects the dynamic interaction between utilization processes and the perceived value of measurement. It can be used to show that priority setting is related to the value that political actors associate with measurement. Two main assumptions must be noted though: first, the level of support can be measured by demand and that continuous demand indicates more support while low level demand indicates the opposite; second, a positive perception will boost and increase demand for measurement while negative perception will suppress or lower demand. In other words, as the perceived value changes over time so is the interest and support of political actors. The model might be used to reflect that when political actors first place demand for measurement it is based on their understanding of the value of measurement. There is an early conception of that value which could be based on prior theoretical knowledge or practical experience. This conception of value determines the purpose of measurement according to stakeholders- multiplicity of stakeholders signifies multiple perspectives about the purpose and the value of measurement. Regardless whether the value is instrumental, performance measurement for improvement (or symbolic performance measurement to comply with institutional pressure) the use is subject to change following demand. And, since supply is the first process that reflects the attempt to respond to demand, stakeholders' perception of measurement might change during or after the processes of supply. When attempting to develop PMS political actors are expected to face political, technical, and cultural challenges that determine the success or failure of adoption. These challenges and subsequent outcomes of the supply process affect the perceived value in the positive or negative direction. This new perception of value affects the demand process and indicates that stakeholders might be either more receptive to or more fearful of measurement. In the same way, using performance information for decision making might satisfy or dissatisfy stakeholders. As outlined by Cavalluzzo and Ittner (2004) and Torres, Pina and Yetano (2011), if stakeholders believe that the new system can support their decision-making activities they are more likely to adopt and use measures, and vice versa. In both cases, supply and use, positive experience enhances the value of measurement leading to more interest, more support and more demand for measurement information from stakeholders while negative experience will likely lead to the opposite. A feedback loop is used to demonstrate how the perception of value changes based on the supply and use processes. It suggests that there is a need to continuously

monitor the feedback loop from supply and use processes to the perceived value in order to anticipate the political mode beforehand and consequently provide the necessary action to maintain or even increase the level of political support for change.

Organizational Capacity

Further factors that affect the utilization of PMS are related to the capacity of public organizations. Organizational capacity in this model refers to the technical aspects that determine whether an organization is capable to provide other non-political support for successful utilization. While the model proposes that organizational capacity is most relevant to the supply process, it also recognizes that organizational capacity can be used to determine organizational readiness for implementation. As suggested by Julnes and Holzer (2001), adopting new systems with disregard to the ability of an organization to implement change might be one reason for failure. Consequently, the model is with the view that while organizational capacity might not guarantee the success of reforms the lack of adequate capacity will likely lead to failure (Moynihan 2004). Reflecting on the systematic review and the empirical work, the model shows that the three main elements that differentiate high- capacity from low-capacity organization are: the level of clarity of objectives and strategies, the availability of sufficient finance needed for operations, and capable workforce that are able to measure and analyze performance data.

Several theoretical and empirical studies suggest that goal clarity is important for successful adoption and implementation of PMS (Cavalluzzo and Ittner, 2004; Bianchi and Rivenbark, 2012; Goh, 2012; Gao, 2015; Jackson, 2005; Pollanen, 2005; Hawke, 2012; Verbeeten, 2008). On the supply side, an organization is more likely to have meaningful quantitative indicators and targets if it is able to clarify and communicate objectives (Hawke, 2012; Moynihan and Pandey, 2005). Conversely, public organizations will find it more difficult to design measurement systems for unclear or vague goals. Difficulties in selecting and interpreting performance metrics negatively impact the development of PM in public organizations (Cavalluzzo and Ittner, 2004). On the use side, ambiguous or vague goals can lead to dysfunctional behaviours. According to Hawke (2012), ambiguity can result in gaming and misrepresentation. It means that there are multiple ways to report performance and how goals are achieved. Internal management may use ambiguity to present information that serves their interest by showing only 'good performance' reports. Further, unclear objectives mean unclear goals of the measurement system which lead to resistance to its implementation by organizational members (De Waal and Counet, 2009). Both the case findings and the literature suggest that organizations' ability to define and clarify their objectives is important for adoption as it is for implementation.

The literature indicates that the availability of adequate financial resources is essential to encourage the development and use of PMS efforts. It points that adequate finance is

essential to improve qualifications and motivate workers. While this could not be proven empirically, due to limited experience in the field of measurement within the study's context, the significance of this factor was discussed in the empirical work paper by pointing to its influence on training and rewards policy. Training is needed to develop the technical skills of those responsible to design measurement systems as much as those responsible for collecting information. Proper guidance is also needed for people receiving and using measurement information (Hawke, 2012). However, the empirical findings suggest that without enough financial resources few public managers will bother to design and conduct training programs that fit the needs of their organizations. Similarly, financial problems could limit motivation by its influence on rewards. Inadequate finance means that organizations will find it harder to reward or celebrate good performance which affects motivation of workers-the discussion about the relevance between the two sub-factors will be presented in a later section.

What's more, the empirical findings and the literature indicate that presence of an effective workforce can facilitate adoption and implementation of PMS. The model consequently suggests that both processes supply and use will need skilled workers who are able to determine what to measure, capable of measurement, and know how to use measurement information for the benefits of their organization. The significance of the availability of adequate skills and qualifications for the supply and use processes merits its inclusion as a standalone factor in the proposed model. However, the case study supports many prior researches in that many government units lack the analytical capacity for PM and that public executives are less familiar with technical aspects of measurement and its functions. The quality of guidance and training is therefore essential for improving the capabilities of workers to understand, interpret, and apply performance data.

Organizational Culture

The proposed model suggests that utilization of PMS in the public sector is also affected by cultural factors. The last decades has seen growing interest on to the relevance of culture and its influence, but there is still limited knowledge in terms of knowing what or how cultural factors affect the implementation of PMS in the public domain (Ohemeng, 2009; Carlucci, Schiuma and Sole, 2015). The performance measurement literature indicates that an organization with a supportive culture for change is more likely to adopt PMS and better equipped to implement the new system more effectively than others (Bititci et al., 2006; Carlucci, Schiuma and Sole, 2015; Taylor, 2011b; Sole and Schiuma, 2010; Sole, 2009; Ohemeng, 2009; Lee and Kim, 2012; Julnes and Holzer, 2001; Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006; Hawke, 2012; Goh, 2012; Gao, 2015; Jankulović and Škorić, 2013). A supportive organizational culture indicates that organizational members understand the value of the new system and accept it as part of their own culture (Ohemeng, 2011; Moynihan and Pandey, 2010). One problem with this definition is that it does not account for factors that shape the views of organizational members and determine their choice(s). Building on the synthesis outlined in the systematic review and reflecting on the case analysis, this model identifies some of the main factors that represent organizational culture and can play central role in the adoption and implementation of PMS in public organizations. The model also proposes that these cultural elements are more relevant to the use process than they are to the supply process. It suggests that the relevance of organizational culture during supply is probably best considered in terms of the influence of certain cultural attributes that could either mitigate or increase resistance to change. The proposed model builds on the work of Carlucci, Schiuma and Sole (2015) to identify the following key cultural factors that affect the utilization of PMS: focus on end results, openness for change and learning, and rewards and incentives' policies. These are now discussed in turn.

Organizational focus on end results is related to the extent to which an organization is guided by its strategies and objectives. Following the guidance of Julnes and Holzer (2001), this could be measured by the extent programs and departments are guided by goals and the extent to which programs and departments communicate these goals. The significance of this factor comes from a view that an organization that is concerned about the end results is keen to improve performance and enhance the quality of decision making. It is also based on the assumption that internal stakeholders who have a normal obligation to produce detailed reports about achievements against pre-determined objectives will see measurement as an addition to their work. They will be interested not only in the development of measures but also keen to use measurement data to improve performance.

Openness to change and learning or 'attitude', as outlined by Carlucci, Schiuma and Sole (2015) and Julnes and Holzer (2001), indicates that organizational members seek out new ideas and want to learn better ways to improve. As delineated in the systematic

review, an organization with a culture that promotes innovation and welcomes new ideas is more likely to adopt and implement PMS effectively. This is because organizations like that are normally in continuous pursuit of learning from their experience and others in order to improve, hence do not see change as a threat but rather as an opportunity. Stakeholders in these organizations would be able to recognize the potential value of the new system and would make a conscious decision to embrace and use it. However, as should be noted from the case study, fear of change and subsequent resistant to it is still one of the main obstacles for adoption and implementation of PMS in the public sector chiefly in developing countries. What can be inferred from the empirical research is that stakeholders, internal stakeholders in particular, are: happier with current organizational practices and norms, not able to see the value of change, and consequently reluctant to adopt new approaches that disrupt normal routine. Accordingly, this model suggests that developing and implementing PMS will be far more difficult without a suitable organizational learning mechanism and a supportive culture for change.

The third factor is related to the availability of rewards and incentive in public organizations. Many empirical and theoretical works suggest that rewards and incentive schemes are essential tools for motivation and for performance improvement that can be used to facilitate adoption and implementation of PMS (Rantanen et al., 2007; Sanger, 2008; Julnes and Holzer, 2001; Lee and Cho, 2011; Lilian Chan, 2004; Sharma and Gadenne, 2011; Pereira and Melão, 2012; Bevanda, Sinković and Currie, 2011; Carlucci, Schiuma and Sole, 2015). These papers maintain that incentives and rewards can be used to boost the use of measures by linking them to performance. Policy of rewards and incentives aims at changing behaviours in the positive direction. It is used as a catalyst to mitigate resistance and gain support for change during the adoption phase. Similarly, it is used as a tool to facilitate the use of performance information for decision making during implementation. This is based on the assumption that establishing a clear link between performance and pay will motivate organizational members and allow them to focus on outcome information less than rules and regulations. Organizational members familiar with being rewarded for good performance will appreciate the value of measurement more than those who are not. While the case findings suggested that it is not possible to apply and use rewards and incentives at all times due to financial constraints in the public sector, it also indicates that most managers consider that successful utilization should benefit from instituting clear policy for rewards.

Interrelation between Processes and Factors

This section intends to describe the interrelations between various elements in the proposed model Figure 19. Suddaby (2010) argues that whilst strong theory building should be based on a clear definition of constructs and their scope conditions, the clarity of constructs is not solely dependent on the precision of the definitions. Understanding and describing the complex relationships that exist between the constructs is important too. Accordingly (notwithstanding the influence of the factors discussed above on the processes of utilization) the model proposes that these processes are influenced by each other and that there is an interrelationship between factors.

Interrelation between processes

The model suggests that demand, supply and use all are interrelated processes that are influenced by each other and by the perceived benefits of measurement. The case study supports the view that performance disclosure in the public sector is mostly driven by demand (Tillema, Mimba and Helden, 2010). Demand might represent an internal need for improvement or compliance with external pressure. The conventional position is that more demand will automatically lead to more supply. As indicated in chapter 5, legal mandates to develop and use PM is one form of demand that have more often lead to the development of measures in various countries. Further, the fact that more supply can lead to more demand was based on understanding how the production process of measures (the supply) affects the perceived value of measurement on the part of stakeholders. The feedback loop between the supply process and perceived value shows that if the process of supply yields a positive outcome, (that is to say, if it is appreciated by stakeholders) those stakeholders will likely demand more performance information based on measurement.

In the same way, the interconnection between ‘supply’ and ‘use’ processes indicates that the supply of measurement data will lead to the actual use of that data for making decisions in public organizations. In line with the view of Julnes and Holzer (2001), the model suggests that the awareness and culture of measurement created by the adoption phase may improve the chances of implementation. Again, this will be related to how the process of supply might have changed the way internal and external stakeholders think about measurement. The literature indicates that the use of measurement data for decision making is mostly determined by the capacity of that data to meet the information need of stakeholders (Taylor, 2011b). When key stakeholders realize the potential benefits of measurement they will be more inclined to use measurement information to take decisions. The quality of the measurement design and the clarity of

purpose are some key elements that affect the perception of stakeholders at this level (Taylor, 2011a).

Additionally, the relation between demand and use is represented by two way arrows which indicates that continuous demand for measurement information will promote the use of that information and vice versa. Increasing the demand for measurement information increases the chances that the information will be used in decision making (Gao, 2015). Continuous demand indicates that measurement is not a symbolic practice that will wane with time but rather instrumental for performance improvement. It represents one form of institutionalization of performance information for decision making purposes. On the same principle, the increase use of measurement information to make decisions most likely boost demand as it signifies the usefulness of PM (Van Dooren, 2005). A feedback link between actual use and perceived benefits demonstrates how the 'perceived benefits' of measurement is affected by the actual use of measurement data. As suggested by Folz, Abdelrazek and Chung (2009), when stakeholders understand the value of measurement information for making decisions they will be motivated to use more types of measures.

Interrelation between factors

According to the literature, there is limited theory on the interaction between factors that affect the utilization of PMS in the public sector (Cavalluzzo and Ittner, 2004). This model endeavours to address this gap by building on the findings from the systematic review and the case study to present some of the interrelationship between factors and the significance of their interaction on the utilization processes.

With regards to political influence on organizational capacity, the model shows that internal and external stakeholders make important decisions regarding key issues such as finance, human resources, and goals. However, the significance of stakeholders' influence might vary according to their organizational role and level of control. Based on the case evidence, external interest groups might have some control over finance and hiring policies but not necessarily on goals' clarity. Consequently, the model suggests that external stakeholders can either promote or hinder the supply process indirectly through their influence on resources. On the other hand, internal stakeholders' influence is clear on all elements of the organizational capacity. While internal management should ensure that the organization has the adequate finance and qualifications to implement programs, it is also responsible for identification of objectives and strategies and to ensure that all departments understand their goals carefully and concisely. Technical problems related to the vagueness of objectives in the public sector reflect the inability of management to break them down into clear and measureable targets which negatively affects the supply and use processes. Without such involvement an organization will likely be hindered by limited capacity which undermines the supply and consequently the use processes.

The influence of internal politics on organizational culture can be better understood by considering the work of Moynihan, Pandey and Wright (2011) who drew upon the trans-formational leadership literature to develop a better conceptualization of the relevance of leadership to the use of measurement information. According to authors, transformational leadership fosters the use of performance data by improving organizational resources and fostering a supportive culture. Rooted in the empirical findings and the literature, the involvement of internal actors in the day to day activities of public organizations indicate that they can choose to make their organizations more open to change and learning, steer the focus on end results, and make decisions to reward or sanction good or bad performance respectively. However, except for their influence on resources, neither the literature nor the research findings provided clear indication concerning the influence of external stakeholders on organizational cultural attributes. Therefore, the model presents the link between political factors and organizational culture in relation to internal stakeholders but not to external stakeholders.

The model also reflects the link between elements of organizational capacity and organizational culture. For example, organizational cultural elements such as the focus on end results and the use of rewards are dependent on the clarity of goals. The latter factor is the basis by which organizational members can focus their attention to the end results. Focusing on end results indicate that an organization is interested in assessing their performance against predetermined goals. Findings support the literature in that having abstract or vague goals will have a negative impact on accountability and on organization's ability to focus on the final outcome (Gao, 2009; Lee and Cho, 2011; Rantanen et al., 2007; De Waal and Counet, 2009). Similarly, considering that rewards and incentives policies intend to motivate staff to enhance their performance and allow organizations to have better control on performance by focusing more on outcome and less on rules and regulations (Lee and Cho, 2011), performance on the other hand is measured against pre-established goals and strategies. The model supports the conclusion of Sharma and Gadenne (2011) in that it is difficult to link rewards and incentives with performance when performance was difficult to measure in the first place. Hence, establishing the link between performance and rewards is not conceivable without clear interpretation of organizational goals.

Likewise, the model shows that rewards and incentive schemes are affected by the availability or non-availability of finance. Recalling that one of the main principles of rewards and incentive scheme is linking pay to performance, findings show that this is difficult if organizations have financial hardships. The data analysis shows that inadequate finance could threaten rewards and incentive schemes which means less motivated workforce. What's more, the model shows that financial capabilities could also impact qualifications, openness to change and learning, and the perceived value of measurement indirectly through their influence on training. Training is associated with building a strong learning and evaluative culture that is open for change and promotes the share of knowledge (Cavalluzzo and Ittner, 2004). It can be designed to reflect and

introduce not only new ideas and technologies but equally important how to deal and manage change. Training can also be used to enhance the perceived benefits in the positive direction. However, my case data shows that because of limited budget in the public sector most managers do not consider training as one of their priorities. The result in this case is that the workforce does not have the knowledge and analytical capacity needed for adoption and implementation of PMS in the public sector which is supported by Yang and Hsieh (2007).

A final remark is related to the type of interaction between various components of the model, be that in terms of relation between processes, between factors or the relations between factors and processes.

For instance, the model notes that the interplay between factors can be viewed as bi-directional where one factor influence and get influenced by another factor(s). For instance, while political actors can have some control over organizational resources they are equally affected and influenced by them. In the public sector, new management doesn't necessarily mean more money or new staff. This means that political actors, management in particular, would be limited by the de-facto capacity and eventually need to adapt and learn how to utilize existing resources for the benefits of the organization.

Another example is when internal actors need to negotiate with and get the approval of external stakeholders regarding resources or certain policies. The ability of internal interest groups to perform their duties depends on a large part on their ability to gain the external stakeholders' trust and support. Less support might mean less finance which undermines the ability of internal management to perform organizational functions.

The bi-directional interplay between the process of supply and organizational capacity is another important point to make in this context. While the discussion above mentioned how organizational resources influence the process of supply, it did not indicate how the latter might influence the former. With that in mind, the model builds on the view of Julnes and Holzer (2001) and Moynihan (2004) in arguing that the interest in PMS could have a positive impact on organizational goals and strategies. Developing certain metrics such as inputs, outputs and outcomes would require goals to be more specific and measurable. Consequently there is some evidence from this study to suggest that the technical nature of measurement could force organizational members to re-question and re-define organizational goals and strategies.

Summary and Conclusions

In this paper I have presented and subsequently explained my model for the utilization of PMS in the public sector. The theoretical model extends the existing literature related to the adoption and use of performance measurement by providing an extensive description about the main factors affecting adoption and implementation of PMS in public organizations. The above discussion indicated that adoption is driven by political and rational considerations whereas implementation is more influenced by cultural factors and the perception of value. Three main differences distinguish this conceptual model from past theoretical paradigms including that of Julnes and Holzer (2001) and Carlucci, Schiuma and Sole (2015).

The first distinction, which constitutes the building block of my conceptual model, is about understanding the meaning of utilization. The model suggests that instead of looking into utilization only in terms of two distinct stages- adoption and implementation- utilization should be viewed in terms of processes at each level of utilization. The above discussion indicates that a process view avoid the cofounding effects of using different terms for similar phenomena and recognizes the significance of political influence by making a clear distinction between real interest in measurement, the development of measures, and actual use of measurement information. Accordingly, whereas adoption is viewed in terms of demand and supply processes, implementation is viewed in terms of the use process.

Second, the conceptual model adds the perceived value as an important element for consideration. It proposes that the perceived value of measurement is a reflection of internal or external political interest in measurement that affects all three processes of utilization. When internal or external political groups realize the potential benefits of measurement they will be more motivated to utilize measurement - they will likely demand, supply, and use more performance measures. Conversely, the model proposes that lack of understanding or appreciation of the value of measurement amongst stakeholders undermines demand, supply and use of measures.

Third, the model does not stop at listing the processes of utilization and the set of factors affecting each but equally important describe the interrelation between all of its components. It consequently proposes that the utilization processes are influenced by each other and that there is an interrelationship between factors. For example, the model proposes that more demand for performance measurement will likely lead to more supply and vice versa- more supply can lead to more demand. On the same principle the model suggests that the awareness and culture of measurement created by the supply process may improve the chances of actual use of measures. Additionally, the actual use of measures may boost demand for more measurement particularly when stakeholders' perception of value is positive. With regards to the key interrelation between factors, the model starts by emphasizing the role of politics on shaping and determining

organizational culture and organizational capacity. However, given the power disparity between stakeholders as outlined by this study, the conceptual model proposes the need to consider the influence of different political groups inside and outside public organizations. The link between elements of organizational capacity and organizational culture suggests the type of interaction between the two set of factors. It basically indicates that a supportive organizational culture is based on the availability of adequate organizational capacity as much it is on the support of political groups.

The presented model has significant value to practice and theory. From a practical point of view, the conceptual model suggests that utilization can be influenced by taking into consideration the following. Managers should be aware about the differences as well as the interrelation between processes and levels of utilization. They should continue to monitor the perceived value of measurement and maintain a positive experience from measurement in order to enhance further development and use of measures. Managers should also be aware about the power imbalance between stakeholders. Initiating change- policy adoption- could be championed by more powerful stakeholders, but actual implementation requires the involvement and support of internal actors. Implementation is also dependent on the provision of needed resources and maintaining a supportive culture. From a theoretical point of view, the conceptual model can be helpful to researchers who are keen to study the use of measurement as a reform tool within public organizations. It stimulates thinking about the main elements that affect successful utilization of measurement and provide a base for discussion about the relevance of possible interdependence between processes and factors to the success of utilization. It calls for shifting the interest of researchers from seeking new factors that affect utilization into questioning the issue of utilization itself and to consider the interrelation between factors and between processes.

However, as this paper indicates, the endeavour to introduce a suitable paradigm to better understand the adoption and implementation of PMS in the public sector is far from over. One should remember that this model is derived from findings of a single study and yet to be interpreted in the context of a larger body of evidence (Rousseau, Manning and Denyer, 2008). The model would need further theoretical testing by more empirical studies to provide guidance about the relevance of all of its elements and their influence on the utilization process. Bearing in mind that public organizations differ in many aspects, some adjustment might be needed according to the specific circumstances of the organization. Similarly, given the noted differences between developing and developed countries in terms of PMS application, assuming that this theoretical model is applicable to all public organizations in all countries is an oversimplification that is not claimed by this study. Further comparative studies that put the assumptions of this model into test should be helpful for researchers and policy makers interested in the utilization issue.

CHAPTER 7: KEY FINDINGS, CONTRIBUTIONS AND LIMITATIONS

Key Findings

The main findings of this DBA are derived from two key studies: a systematic review and an empirical research. The first project was interested in forming an understanding of the main elements that determine the success of PMS in public organizations. The second project was more interested in exploring the experience of PMS in the public domain in non-western countries by looking into the patterns of measures use in several organizations in the Yemeni public sector. The empirical project built on the first project to look into what factors affect utilization in practice within the context of developing countries and possible resemblance and dissimilarities with that in developed countries. The following section provides further description about the main findings from the literature and from empirical project.

Key Findings of Project 1

The findings of the first project provided important information regarding the definition and identification of the elements of PMS utilization in the public sector. The study supported the view of Julnes and Holzer (2001) in defining utilization of PMS as a process that consists of two main stages adoption and implementation.

Utilization Processes

In line with Yang and Hsieh (2007), the Systematic Review shows that most past studies have been looking at determinants of successful utilization of measurement, often without explicitly defining the meaning or the boundaries of the utilization process which adds to the complexity of research in this field and inhibits its development. It also could have serious implications on determining the factors that affect each stage. For example, the Systematic Review paper notice that ‘implementation’ is not the same for all papers; some papers use it to refer to adoption while some others use it to refer to the actual use of measures. To avoid confusion about the exact meaning of different terms, determinants of successful utilization were examined in terms of their relation to the supply and consumption of performance information. Adoption was examined in terms of demand and supply of performance information, and actual use in terms of making decisions based on measurement data.

Size and Interest on the PMS Field

The Systematic Review revealed that the number of empirical studies on PM is still relatively small compared to the overall literature in this field. It also shows that while there is growing interest in the field of PM many of the study published has been confined to western and developed world. The review suggests that PMS has been an

accepted practice that is taken for granted in these countries. Consequently, the interest of key stakeholders in that context is focused on learning the best way to develop and use measurement efficiently and effectively.

Main Drivers

The Systematic Review indicates that PMS adoption is first and foremost about the demand of performance information from the part of stakeholders. It shows that the multiplicity of stakeholders in public organizations signifies multiplicity of drivers of adoption as much as for expectations. The review indicates that internal and external actors are motivated by the value they associate with measurement initiatives. When internal or external stakeholders realize the potential benefits of the measurement system they will put pressure/demand towards its adoption. Accordingly, positive perception will initiate/increase demand and negative perception will lead to the opposite- blocking/suppressing demand.

Utilization Factors

The Systematic Review reflected that papers in the review had different perspectives in terms of 'what' and 'how' certain elements determine success or failure of PMS in the public sector. While main elements that affect adoption and implementation have been noted by most papers, there was clearly a lack of consensus on specific class or categorization of these factors. In fact, most papers avoid the use of classification of factors basically listing factors relevant to their context. For example, the majority of papers simply mentioned that leadership support, involvement of stakeholders, availability of adequate resources and training is important for utilization. Upon reflection on the literature, the Systematic Review paper concluded that adoption and implementation are affected by the same group of factors albeit the level of influence and relevance of each group of factors will vary at each stage of utilization. Accordingly, I decided to group the main factors that affect utilization into: political, technical, and cultural factors.

Political factors encompass the influence of all stakeholders inside and outside organizations. Internal politics according to the majority of papers in the literature is related to the role of management and employees, while external politics, include but not limited to, the role of elected officials, regulators, work unions, and citizens. The Systematic Review indicates that internal desire for improvement and external pressure are the main driving factors for adopting PMS in the public sector. It indicates that internal support and leadership is crucial for the development of measures as it is for integrating measures for the purpose of decision making. However, the Systematic Review provided little account about the influence of external stakeholders on the actual

development of measures implying that external politics have a partial influence on utilization by creating the desire for change and not necessarily affecting change itself.

Technical factors that were mentioned by most papers in the review related to the available capacity for an organization to develop and use PMs. It includes elements such as finance, qualifications, training, and measurement design. Technical factors were grouped into three main categories: measurement issues, resources' availability, and the design of the measurement system. With regards to the first category, many papers suggested that there are many problems in relation to measurability of public services due to the difficulty to measure performance in the public sector. As noted by many, one problem is that goals in the public sector are vague, broad and unachievable. Further, the review indicates that considerable resources (financial and human resources in particular) are needed during and after the development of measures. Similarly, the quality of the technical design of the measurement system is crucial for adoption and implementation. One fundamental basis of any design for measurement is its ability to define measures or indicators concurrent with the vision and strategic objectives of the organization. However, the literature reveals that there is a deficiency in terms of providing public organizations with measurement systems that are tailored to their needs. Instead, the public sector continues to use and apply different practices and frameworks that are developed for and used by the private sector.

Cultural factors are related to all elements that define and shape the culture of public organizations and have an impact on adoption and implementation. Based on an initial analysis of the literature (only few papers were interested in cultural influence), I decided that the basis of cultural influence consists of two main determinants: performance-oriented culture and openness to change and learning. A performance oriented culture emphasizes that performance measurement is viewed as a vehicle for improvement. Openness to change and learning indicates that organizations are in continuous pursuit to learn from past experience and search for better ways to improve performance. Accordingly, the literature suggested that organizations that have those cultural determinants have better chances to develop and use measurement successfully. Other cultural considerations in public organizations in non-western countries were also noted in the review. Problems of informality, low-institutional capacity and corruption are some elements that might undermine the potential for developing and using PM successfully in public organizations in developing countries.

From Measurement to Management of Performance

The literature indicates that PM has been an accepted practice in most countries around the world. The potentials of measurement to improve efficiency, minimize waste, and provide value for money services is widely recognized and accepted by stakeholders in the public sector and no longer a matter for challenge. Consequently, there is a shift from questioning the advantages of PM to questioning the elements that determine successful adoption. However, the literature also shows that the development of

measures doesn't necessarily mean that PMS will automatically be integrated into the culture of public organizations. Despite their potential to improve performance, many public organizations do not use measurement to take decisions. This could explain the growing interest to learn more about the link between measurement and decision making and the need to identify the factors that affect the relationship between the two. Accordingly, the existing literature suggested a shift in discussion from performance measurement, which is merely interested in quantifying and determining performance of an organization or program, to the broader approach of performance management that is concerned with using measures to enhance decision making and the overall performance in the public sector. Additional general findings and gaps in the field is presented in the following table.

General Conclusions from the Systematic Review
Empirical work remains thin and focused on single issues
No conclusive distinction between different levels of utilization
Lack of an integrative perspective that determine the influence of various factors on each level of utilization
Deficiency of knowledge in relation to how factors interrelate and influence each other
No consideration for the dynamism of political support and the elements that boost or undermine that support
Most published papers are in English language mostly covering western countries with little attention to other countries
Utilization of PM in the public sector is work in progress with much to learn from the experience of various public organizations in different contexts

Key Findings of Empirical Project

The findings of the empirical project provided an understanding of key issues related to the development and use of PMS in public organizations in non-western countries. Given the limited research on public reforms in developing countries, there is clear deficiency in the ability to understand whether PM serves similar functions in that context as it does in the western world and to grasp whether adoption and implementation of PM faces the same institutional and capacity constraints regardless of the context (Gao, 2009; Ohemeng, 2009; Holmes, Amin Gutiérrez de Piñeres and Douglas Kiel, 2006; Tillema, Mimba and Helden, 2010). By examining the views of several public officials on the actual and potential role of performance measurement to support reform and improve performance in the Yemeni public sector, the project provided useful information about the arrangements that promote or impede the development and use of PMS and subsequent importance of the context. Consequently, assuming that PM design and implementation are faced with the same key challenges in all public organizations in all countries Abubakar, Saidin and Ahmi (2016) is oversimplified and unwarranted claim that is challenged by findings of this project.

The empirical project revealed the following key findings:

Public Administration

My reflection on the interviews indicates that the principles of NPM governing the western world are not necessarily copied in developing countries. The Yemeni public sector features a traditional administration that is far from being a result-oriented management, bond to roles and procedures, and less flexible to change. Consecutive reforms in the past years did not incorporate a clear policy for performance measurement and management. Performance is mostly determined by classical methods based on subjective evaluation rather than actual measurement of performance.

Cognitive Barrier

Managers' answers indicated that there is limited knowledge and understanding of the use and usefulness of measurement amongst the majority of stakeholders in the public sector. The study reveals that cognitive limitation might be a prime limitation for the adoption and implementation of PMS in public organizations in the context of developing countries. Cognitive limitation suggests that stakeholders are unable to determine the purpose and use of measurement to derive up performance. Lack of knowledge about the value of measurement, lack of skills to design proper measurement systems, and lack of experience in this field are all results of that limitation. Consequently, the study suggests that institutionalizing principles of performance

measurement and management through legislations might be helpful for initiating change in the public sector, particularly in developing countries.

PI Demand and Supply

Findings suggest that there is insufficient demand for performance information in public organizations which effectively led to limited supply. Public organizations are dominated by a weak accountability culture that does not promote the communication of performance data from one level of government to another. The interviews revealed that there is hardly any accountability for results and that present legislations and government mandates do not seem to increase demand and supply of performance information. This revelation indicates that adopting PMS in that context is limited by insufficient demand for and supply of performance information.

External pressure

The research shows that performance disclosure in developing countries is mostly driven by external rather than internal factors. Internal stakeholders may have some reservations in terms of revealing unnecessary information about performance that could either harm them or cut down organizational resources. Taking into consideration stakeholders' limited perception of the value of measurement and weak accountability culture, the research suggests that adopting PMS is less likely to be the choice/decision of internal actors. External requirements and pressure on the other hand can be one suitable driver for the introduction of PMS in public organizations in the context of developing countries. The study support the findings of Mimba, Helden and Tillema (2013) that funding bodies are the dominant stakeholders with the strongest power position to induce change. Having said that, results also suggest that external pressure alone is not sufficient for success. The research maintain that legislations and external mandates can create the first step (adoption) but the actual use and subsequent value of measurement is less certain and dependant on additional factors.

Managerial Discretion

The research noted that many important financial and employment decisions are made by external stakeholders. There is limited discretion given to management in terms of control over resources in many public organizations which undermine their motive to support change. For example, managers find it difficult to reward good performance financially because of their limited control on budget. The study concludes that managers who lack the authority to make decisions based on performance have little incentive to support measurement initiatives. However, managers have more freedom in terms of reporting performance. It could be that the absence of external scrutiny made it possible for some managers to reflect their own success and hide failure.

Main Barriers

Certain challenges related to organizational capacity and dominant culture in developing countries that limit public organizations' endeavour to adopt and use PM in the public sector were also noted by the study.

Fiscal constraints: this is the first cited factor by the majority of interviewees which is mostly related to limited budget and improper financial policies. Financial constraints according to interviewees undermine any attempt for change in the public sector. Linking performance to pay and training are two important functions that can't be implemented due to financial problems. However, the research also notes that not all public organizations face the same level of financial hardships. Other organizations are less dependent on government's support and have other means of finance.

Limited qualifications: The study shows that organizational members: are unable to understand and determine the purpose of measurement; do not know how to measure performance; unable to interpret performance data; and consequently unable to use performance information to take decisions. The study suggested that training is essential in order to improve the cognition about the need and value of measurement, cover the design of proper system and selection of measures, define goals and targets, and to communicate performance expectations.

Unclear objectives: most public organizations have vague and unclear objectives and there have been limited attempts to address this matter in the Yemeni government. However, the paper also notes that public organizations with more tangible outputs and standardized services, such as line agencies, are better in defining some of their goals than others and consequently better able to assess their performance.

Resistant: fear of change and subsequent resistance to it may be a fundamental obstacle to the development and use of measures. Internal stakeholders are normally not risk takers and feel more comfortable with traditional practices and norms. They need to see the value of change and view it as an opportunity rather than a risk.

Weak incentive schemes: managers agreed that for PMS to work in public organizations there should be a strong incentive system that rewards and sanctions good and bad performance respectively. However, many of which admitted that they were unable to enforce the incentive scheme due to some limitations chiefly related to their inability to link individual work to the overall performance and lack of adequate finance.

Weak accountability: reporting is a symbolic practice in most organizations. There are limited guidelines about how to report performance and lack of feedback on these reports from upper level government. Weak accountability signifies that an organization is not interested in measuring its performance. The research concluded that weak

accountability could be more relevant factor than corruption which is difficult to measure.

Organizational Differences

The study notes that the public sector is large and complex consisting of various types of institutions that varies in many aspects, such as services, goals, functions, structure, finance, and management. Three main categories were used to describe the Yemeni public sector: (a) central agencies (b) line agencies and (c) state-owned enterprises. The research concluded that a one-size-fits-all policy should be avoided by policy makers who should be mindful of possible dissimilarities between public entities. It demonstrates that one possible reason for the lack of influence of external mandates on the implementation of measurement is that such policies often disregard an organization's ability to implement change.

Interdependence between Factors

The assumption that various political, technical, and cultural factors independently influence utilization is challenged by the research findings. The study suggests that factors do interrelate with each other and provided some examples of that interaction. Consequently, the study adds to the voice that there is a need to recognize the cause and effect mechanism of the interaction between factors in order to understand how certain problems happen and why and in order to determine and isolate generic problems from those limited to their own context. It also supports the need to determine the relevance of factors possibly by distinguishing between important, promising and insignificant variables (Kroll, 2016).

Contributions

Contribution to Knowledge

This study extends the current literature related to the adoption and implementation of PMS in the public sector in general but with special attention to the context of developing countries. Its findings suggest that failure and success of utilization of PMS in the public sector is influenced by various elements that need to be considered prior and after the adoption of measurement. The following are the main theoretical contributions of this study.

1. The proposed conceptual model of the utilization of PMS in public organizations represents the most noteworthy contribution of this study. The model provides a comprehensive view about the main elements that affect adoption and implementation of PMS, and how. The value of the models' contribution is three-fold. First, the model can be used as a guideline for identifying and exploring significant components that constitute and affect utilization of PMS in the public sector. It promotes understanding of the main elements that affect successful design and implementation of PMS in the public sector which has not been receiving wide scholarly attention. Second, the model provides the foundation for building and testing additional hypotheses (propositions) on the variables that influence adoption and implementation policies in the public domain in different contexts. It can be used to form generic checklist on key success factors and/or potential challenges in different phases of utilization. Third, the model not only provides a useful overview about the level of significance of each factor at each level of utilization, but equally important provides an explanation of possible interrelation between these factors and how such interaction affects processes of utilization. The model can be widely used for research purposes to form better understanding about the shift from performance measurement to the broader and supplementary approach of performance management in the public sector.
2. This study recognizes that performance measurement and management is not common knowledge to all countries. It shows that the lack of understanding about the benefits and use of measures could be a major barrier that limit or block successful utilization of PMS in non-western countries. Recognizing the lack of cognition about the value and benefits of PMS in the public sector as a possible barrier for adoption in the context of developing countries has not been reported previously to my best knowledge.

3. The study reinforces the importance of recognizing and understanding how political actors perceive the value of measurement. Results suggest that there is limited research on this issue and that without understanding how stakeholders perceive the value of measurement our knowledge about the utilization of PMS in the public sector would be incomplete. The study **concludes** that adoption and implementation of PMS in the public sector is directly related to how public officials perceive the use and usefulness of measures.

4. Further, the study **adds** to the body of knowledge by shedding the light on the dynamic of management commitment to the development and use of performance measurement in the public domain which received little attention in the literature. It indicates that stakeholders' perception of value is changing and so is their level of support. As stated earlier, while some previous researches show a direct link between adoption and perceived benefits this research show that continuous support of stakeholders is also affected by the value of measurement. Accordingly, the study suggests that there is a dynamic interaction between demand and perceived benefits of measurement. It also **demonstrated** that demand could be one measure for how stakeholders value measurement. More demand for performance information indicates that stakeholders' perception is positive about the value of measurement and low demand indicates that their perception about the value is negative or has changed into the negative direction. The research advocates that maintaining positive perception is important for stability and continuity of political support.

5. This study built on the framework of Julnes and Holzer (2001) in viewing utilization of PMS as a change process involving two main stages adoption and implementation. However, while the study confirms that adoption and implementation are two different stages that influence each other and are not equally determined by the same set of factors, it also suggests the need for translating abstract concepts such as adoption and implementation into clearly defined theoretical constructs in order to avoid confusion and create a common vocabulary for researchers. Consequently, the study suggests a need to make a clear distinction between real interest in measurement, the development of measures, and actual use of performance information. That's one reason that explain why defining PMS utilization in terms of processes is a **substantive contribution** to theory and practice. The study **proposes** that demand can be used as a standalone process to explain and monitor the level of involvement and commitment of stakeholders in the development and use of PM. It shows that adopting PMS is not distinct from the will of political actors inside or outside the organizations and stresses the role of power and politics in the overall process of utilization. The supply process may refer to the actual development of performance measures. It focuses on determining whether an organization has

been able to meet pre-established objectives using measurable standards for performance. The use process is about what organizations decide to do with acquired knowledge about its performance. It is concerned with the ability of stakeholders to make an informed decision(s) based on performance information. Accordingly, while adoption is explained in terms of demand and supply processes, implementation is concerned with the use process. Differentiating between different levels of utilization and looking into adoption and implementation in terms of relevant processes should provide a better picture about the influence of various factors on each and consequently enhance communication amongst the research community.

6. The study reflects the need to understand the limits and boundaries of internal and external political influence. It adds to the body of knowledge by emphasizing the role of power and pointing to the source of change in the public domain. The study confirms the power distance between different stakeholders and concludes that any attempt to adopt PMS in the public sector should take into account different power positions of different stakeholders and study their influence.
7. Whilst this research concurs with the results of previous research in concluding that external political pressure is the main driver for the introduction of PMS in public organizations, results of this research show that external political influence is not sufficient for success as its span of control does not cover all dimensions of organizational capacity and culture. According to findings, external stakeholders have a direct influence on the demand process and indirect influence on supply through their control over resources. Findings also suggest that external stakeholders have limited or no control over the use process. This is **contrary** to the argument provided by some scholars including Carlucci, Schiuma and Sole (2015) and Julnes and Holzer (2001) who emphasizes the positive effect of external politics on implementation. Given the discrepancies in results and research's strategy, this study highlights the need for further empirical research to better understand the influence of external stakeholders on adoption and implementation.
8. As for the role of internal stakeholders, data confirms prior findings that it is important and relevant for both adoption and implementation. Internal actors initiate demand, select and develop performance measures, and choose to use or not use performance information for making decisions. However, the study also **reveals** that internal political influence is largely dependent on the level of discretion given to internal actors- mainly managers- to take decisions based on performance information. Given the limited empirical work and knowledge on this issue the study calls for further research to investigate how managerial discretion can influence both stages of PMS utilization in the public sector.

9. The findings of this research support prior studies in indicating that adopting new systems with disregard to organizational capacity might lead to failure. Upon reflection on the literature and empirical findings, the study **proposes** three main elements that could be used to differentiate high- capacity from low-capacity organization: the level of clarity of objectives and strategies, the availability of sufficient finance needed for operations, and capable workforce that are able to measure and analyze performance data. Several papers investigated the link between goals' clarity and measurement initiatives. According to those papers, specific and clear goals are easy to measure while vague goals are elusive and hard to capture or measure. However, this study shows further significance of goals' clarity on utilization. The study empirically demonstrates that goals are one key element that could be used to distinguish between different types of public organizations. It shows that goals are generally broad and vague in central organizations while they are more specific and tangible in line agencies. Consequently, the study proposes that line agencies will be more inclined to adopt and use PMS than central agencies.
10. In the same way, many papers refer to resources' adequacy as an important technical element for successful development and use of PMS but with no proper definition of what aspects of resources they have in mind - it is not always clear whether it is time, money, human capabilities, or other aspects of resources. This study, on the other hand, provides better insight by making clear distinction between financial and human resources, and by showing the relevance of each on the utilization processes. While it support previous findings on the significance of capabilities on the supply and use processes, it also builds on the empirical findings to propose that finance does not affect the supply and use processes directly per se but rather indirectly through its influence on training and incentive schemes. One implication that needs to be verified by further research is that financial deficiency does not necessarily limit adoption but could limit implementation.
11. The study shows that culture is defined and explained in the literature in numerous ways and verifies that there is limited study about cultural influence on utilization. It joins and adds to the current debate by identifying which cultural aspects are related to utilization of PMS in the public sector and by showing their level of influence at each stage. The study suggests that the three cultural elements- maintaining a culture that is open for change and learning, focusing on end results, and celebrating and rewarding achievements- will mitigate resistance during supply and promote the actual use of performance information for decision making. One advantage from identifying these cultural attributes is the ability to make comparison between different public organizations in multiple contexts. Researchers and policy makers will be better

equipped to focus on cultural elements that promote the development and use of PMS in public organizations.

12. This research shows that there is considerable value in studying the application of PMS in various government agencies. The study confirms that the public sector is large and complex system consisting of many types of organizations that differ in many aspects such as size, structure, finance, management and culture, and provides various types of services and functions. It provides scrutiny of different processes and behaviors in different socioeconomic settings in order to identify common elements from those specific to their context. The study **empirically demonstrates** that differences between various types of organizations could have serious implications on the utilization of PMS. It shows that a one-size-fits-all policy should be avoided by policy makers who should be mindful of possible dissimilarities between public entities. Accordingly, the study suggests that PMS utilization should be attempted in accordance to the organizational capacity and culture of each organization. Having said that, it should be noted that it is difficult to generalize findings from one single research which is why further investigations are needed to confirm the relevance of the research's findings. As outlined by (Bianchi and Rivenbark 2012), comparative case studies are useful to discover how two or more cases compare in specific outcomes, conditions and criteria.
13. One of the **main contributions** from this research is that it provided some explanations about possible interrelations between political, technical and cultural factors and how such relationships affect the utilization process. Investigating the relationships between factors is helpful to gain better understanding about the significance of each factor and to learn more about the differences in the implementation of PMS across various types of organizations. If previous researches were solely focused on factors that affect utilization with no proper consideration of possible inter-relation between them, as suggested by Cavalluzzo and Ittner (2004) and Hawke (2012), this means that further research is needed to better understand the inter-dependence and influence between factors and how addressing basic challenges can mitigate smaller ones.
14. This study has helped to address the knowledge gap that still exists in the literature about the application of PMS in non-western countries. It worked to promote and enhance understanding about the level of similarities or dissimilarities, in terms of PMS utilization, between organizations in different countries. It shows that public organizations in the non-developed world face different challenges and different local circumstances that could influence the characteristics of the public sector. For example, the study **reveals** that public organizations in developing countries are generally characterised by weak accountability culture, low involvement of stakeholders, lack of understanding about the use and usefulness of PMS, limited finance, limited managerial discretion, and weak skills amongst public workforce. The study makes further

contribution by **generating two important prepositions** related to the significance of the context. Recalling that findings suggest that low level demand for performance information and subsequent limited supply can be attributed to the weak accountability culture in public organizations within the context of developing countries, the first preposition is related to the link between accountability and measurement. While the majority of literature focuses on PM as a tool to promote accountability, the study suggests that greater accountability for results should provide the necessary incentive for measurement. It proposes that boosting accountability culture will increase the chances for adoption through its influence on the demand process. Second, given the weak accountability culture in developing countries and subsequent low demand for performance data and lack of understanding about the use and usefulness of PMS, the study suggests that the chance of voluntary adoption of measurement is slim in that context. Accordingly, the study proposes that legalization could be better option for institutionalizing PMS in public organization in non-western countries.

15. In conclusion, the study opposes Abubakar, Saidin and Ahmi (2016) claim about the universality of PMS application regardless of the context. It conversely concludes that there are significant cultural differences between countries that have serious implications on PMS utilization. It confirms the conclusion of Gao (2015) that the context always matters and that each country and organization has its own expectations, formula, strength, weaknesses, and pace of change. Further empirical research should be helpful to confirm or challenge these contentions.

Contribution to Practice

The results of this research provide a number of recommendations that could be helpful for policy makers in Yemen in particular and for practitioners in general to better understand and address some of the challenges and limitations for utilizing measurement for reforms in the public sector. By making comparison between what the literature says about this issue and what key stakeholders consider important, the study puts forward some of the key issues that need to be addressed to utilize measurement effectively in the public sector.

1. Increasing Awareness and Understanding

This study shows that cognitive barrier could be a major concern in developing countries. It reveals that public organizations may not be able to utilize measurement when key actors are not familiar with the concept of measurement and its functions. Accordingly, the study signifies the importance of increasing awareness and understanding of the need for change prior and after change takes its course. Training should be helpful to provide organizational members with the right mechanism to understand, accept, and lead successful utilization of measurement. Policy makers can start by employing training to increase awareness and understanding about the benefits of measurements for the purpose of accountability and performance improvement. It should be viewed as an incentive tool to inspire stakeholders to adopt measurement systems and reduce their resistance to change. Given the noted limited experience in measurement in the Yemeni context, training is also needed to enhance the technical capacity of internal actors in that field. The supply process entails that internal actors should be able to answer the following questions:

- What is the right design for measurement that fit the purpose?
- What to measure? And how?
- How data can be interpreted and communicated?

The conceptual model shows that training might be also useful to reshape organizational culture in a way that makes public organizations more open to innovations and new ideas as well as increasing stakeholders' willingness to learn from past performance. Policy makers can use training to communicate the purpose of the new system in order to address fear from change and subsequent resistant to it. To that end, the study suggests that the provision of resources for

training is necessary and an important indication of the level of seriousness and commitment of key stakeholders for successful utilization.

2. Managerial Role and Leadership

The study signifies the importance of management's role in all processes of utilization. Lack of or limited managerial support and involvement during all processes of utilization could either lead to failure or to the development of a symbolic practice. This means that managers should reflect their interest and commitment during all stages of utilization. They should define the quality of information they require and show their willingness to use that information for decision making and performance improvement. For that to happen, managers should have the expertise, capacity, and power to control resources and make decisions based on performance information. The noted limited authority given to managers to use financial resources to celebrate achievements in the Yemeni public sector could be a major barrier for successful utilization in that context.

3. Monitoring the Value to Stakeholders

This research establishes that political support and commitment is not constant but rather dynamic that continues to change over time. It is linked to the perceived value of measurement and the way that the measurement system can serve the interest of key stakeholders in the public sector. It consequently highlights the need for continuous monitoring of the value of the new system to key stakeholders and for boosting demand for measurement information. A good idea could start by developing a feedback sheet that can be used on regular bases to measure the level of satisfaction of major stakeholders about the usefulness of measurement information. Another suggestion is monitoring demand for performance information which could give an indication about the level of stakeholders' interest in measurement. As indicated earlier in this paper, positive perception will likely lead to more demand for performance information while negative perception will lead to the opposite. Continuous monitoring of the value to key stakeholders provides an opportunity for identifying problems and correcting actions beforehand which ensures the continuous usability and relevance of the new system.

4. Organizational Readiness Assessment

As indicated earlier in this paper, this study recognizes that public organizations are not necessarily similar to each other and that there are considerable differences between different types of organizations. It maintains that conducting an assessment of an organization's resources and culture could be helpful to understand potentials versus limitations in order to determine

organizational readiness for change (Julnes and Holzer, 2001). The main elements that were defined in the conceptual model could be helpful to understand the main challenges for effective development and use of measures. The significant of these elements is that they provide a picture about the level at which these factors are most relevant and how. However, while this study recommends that the scope and speed of change should be adapted according to each organization's resources and culture, it also suggests that incremental adoption of performance measures can enhance the level of awareness about measurement and affect organizational culture towards further adoption and implementation. This simply means that managers and policy makers should understand that the process of utilization is incremental and evolutionary- PMS is not automatically integrated into organizational culture based on a full-scale approach. In the experience of many public organizations in developed countries, the process was rather incremental and took considerable scope of time.

5. Increasing Accountability and Demand for Performance Data

This research's findings show that reporting is a symbolic practice in the Yemeni public sector and reveals the weak accountability culture in that context. Noting the positive relationship between increasing accountability and chances of PMS adoption and use, the study concludes that key stakeholders should work to boost accountability by increasing the demand for performance information and by getting involved in determining the type of data that fit their needs. This applies to both internal and external stakeholders.

6. The Role of External Stakeholders

One central argument of this paper is that change in public organizations in developing countries is less likely to happen without external intervention. The study notes the significant variations in the scope of political influence amongst stakeholders and concludes that donors and funding agencies appear to be the most influential stakeholders that have the power to induce change in that context. Donors and funding agencies can, therefore, push for institutionalize PMS through legislations or administrative directives, or at least use their finance's leverage to press for the introduction of measurement through continuous demand for performance information that is based on actual measurement and not on subjective evaluation. Having said that, it should be noted that while this study acknowledges that external pressure might be helpful for initiating change, possibly leading to the development of measures, it also recognizes that external pressure is not enough on its own to ensure the actual use of measures.

7. *Goals' Orientation*

This study shows that ambiguity of performance objectives will continue to impede an effective development and use of measurement in public organizations. It signifies the need to work on clarifying organizational objectives and strategies in a way that makes them measurable and understandable. Managers of public organizations, particularly in central agencies, are requested to take that into consideration when attempting to develop measures of performance. They should ensure that organizational departments are capable of constructing clearer and achievable targets and be able to communicate them to different levels in the organization.

8. *Need for Resources*

As indicated earlier in this research, financial resources are necessary for improving qualifications and acceptance of measurement through training that is required before and after the development of measures. Financial resources are also needed for schemes of rewards and incentives' policy. Key stakeholders interested in successful utilization of PMS in their organizations should be prepared to provide adequate finance for the new system.

9. *Promoting Result-based Culture*

The study calls for building a culture that focuses on end results and celebrates achievements. Public organizations in developing countries need to shift from traditional mechanism of conducting business into more flexible and result-based management if they wish to succeed in using measures for reform. Focus on end results means that organization would have a clear and understandable goals and strategy. It encourages a deliberate quest for performance information in order to learn how to improve performance. Linking performance to pay might be helpful to ensure that the system is not perceived as symbolic but with actual value to organizational members.

Limitations and Suggestions for Further Research

This section discusses some of the main limitations of this study and its contribution to knowledge and practice. Limitations are explained in terms of generalization of findings, replicating the research, respondents' understanding, and personal bias, paving the way for some suggestions to mitigate problems and guide future research.

Generalization Issues

This paper reports the findings of a qualitative study, based on in-depth interviews with many government officials in the Yemeni public sector. While a qualitative approach allows for further understanding of the issues to be investigated, it is generally based on small sample size which limits the possibility to generalize findings. Similarly, the problem of using non-probability sampling is that there is little evidence that they are representative to the whole population. While a larger and random sample that include participants from different kinds of public organizations may seem ideal for generalization purposes it is difficult to achieve in reality, particularly under similar political and social conditions in the country during the time of the inquiry.

Another element that limits generalization is that this research was conducted on no specific type of public organizations. Hence, whether findings are mirrored across the Yemeni public organizations will only be determined by further researches that include more and divers organizations.

However, as indicated in Chapter 5, this study is less concerned with issue of representativeness, which is less important in qualitative researches (Bryman and Bell, 2007), than understanding the area of the study itself. It follows Yin (2002) in suggesting that while the analytical benefits from having multiple cases may be substantial, findings of this type of research might best be generalized to theoretical propositions.

Replication issues

Another limitation is related to the ability to replicate the research. 'Replicability' is the degree to which the findings of a study can be reproduced (Bryman and Bell, 2007). Replication might be helpful for external validity and generalization purposes. In this research, I worked to be as explicit as possible about the procedure to ensure replicability. However, given the conditions that governed this enquiry, discussed in details in Chapter 3, this paper acknowledges the difficulty to replicate the same research process.

Understanding

Since findings are based on the views of public managers, it is not known to what extent their views reflect reality. While interview questions were set to solicit managers' understanding of performance measurement and management, it is possible that they may have different interpretation of the concept than that in the literature. Accordingly, the sole reliance on the opinions of officials alone may not provide sufficient information about the context and the factors that distinguish between public organizations in relation to the adoption and use of measures.

Research Bias

In-depth interviews, as a research method, require a high level of involvement of the researcher. Ideally, the results of a certain research should be unaffected by the researcher's special characteristics and expectations (Bryman and Bell, 2007). However, it is not possible to know for sure, in my opinion, that the perceptions and believes of the researcher do not affect the design and outcome of the research. For example, the study findings are based on a continual refinement of and reflections on gathered material which could be subjective and biased. However, while there was no clear evidence for lack of objectivity or personal bias, this research does not claim to provide a definitive picture of the world. It rather suggests that the limitations of this research should be overcome with further research that is preferably based on the work of multiple researchers instead of only one.

Opportunities for Further Research

The study shows that further research is needed to understand the conditions that govern the development and use of PM in public organizations, particularly in the developing world. As indicated earlier in this paper, little is known about the organizational and contextual factors that determine the success or failure of PMS utilization in the public domain. The study consequently suggests that there would be considerable value in investigating the same research questions in different contexts on multiple and diverse public organizations. Future empirical studies and concurrent theoretical testing would provide useful guidance to harness the potentials of measurement to reform the public sector and increase its efficiency and effectiveness. This could be helpful to challenge or confirm the foundations of the presented theoretical model. Based on the results of this study, these are some suggestions for issues to consider for further research in the future:

- The study maybe advanced methodologically by adopting multiple research strategies (quantitative and qualitative) on larger sample size representing various types of public organizations.
- A comparative case study between public organizations in different countries could be helpful to determine generic and specific elements that affect utilization of PMS in the public sector.
- Data collection could be enhanced by examining available documents of each case. Sole reliance on the narrative and views of respondents can't provide an accurate picture about the issue under investigation.
- This study suggests that factors do interrelate and influence each other. And, given the limited research on this area, future research should put more emphasis on the relationship between factors in order to understand the influence of each group of factors on each level of utilization. Investigating the link between factors is also helpful to determine generic factors from those specific to their respective context.
- One of the main premises of this study is that the interrelation between processes is an important element to consider for successful utilization. The study suggests that future research should be mindful about the interrelation between process and consequent influence on the utilization issue.

To conclude, this paper suggests that integrating PM into public management is a dynamic process that cannot be based on a single decision or intervention. It calls for further research to update the list of factors that affect PMS utilization in the public sector. While results of this study could be retested and challenged to enhance the statistical power to generalize findings, the body of knowledge could also be enhanced significantly from extending the inquiry to other non-western countries. Findings also indicate that there is a need for improving theorization about utilization of PMS in the public domain. While the conceptual model was based on extensive review of the literature and field study, it is still need to be tested and challenged empirically and theoretically.

REFERENCES

- Abubakar, A., Saidin, S.Z. and Ahmi, A. (2016) 'Adoption of the performance measurement system in the Nigerian public sector: Challenges and prospects', *International Journal of Management Research and Reviews*, 6(10), pp. 1424-1436.
- Adcroft, A. and Willis, R. (2005) 'The (un)intended outcome of public sector performance measurement', *International Journal of Public Sector Management*, 18(5) pp. 386-400.
- Akbar, R., Pilcher, R.A. and Perrin, B. (2015) 'Implementing performance measurement systems: Indonesian local government under pressure', *Qualitative Research Accounting and Management*, 12(1) pp. 3-33.
- Ammons, D.N. and Rivenbark, W.C. (2008) 'Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project', *Public Administration Review*, 68(2) pp. 304-318.
- Bamford, D.R. and Forrester, P.L. (2003) 'Managing planned and emergent change within an operations management environment', *International Journal of Operations and Production Management*, 23(5) pp. 546-564.
- Barton Cunningham, J. and Kempling, J.S. (2009) 'Implementing change in public sector organizations', *Management Decision*, 47(2) pp. 330-344.
- Bartunek, J.M. and Rynes, S.L. (2010) 'The Construction and Contributions of "Implications for Practice": What's in Them and What Might They Offer?', *Academy of Management Learning and Education*, 9(1) pp. 100 -117.
- Behn, R.D. (2003) 'Why Measure Performance? Different Purposes Require Different Measures', *Public Administration Review*, 63(5) pp. 586-606.
- Berg, B.L. (1998) *Qualitative Research Methods for the Social Sciences*. 4th edn. Boston: Allyn and Bacon.
- Berman, E. and Wang, X. (2000) 'Performance Measurement in U.S. Counties: Capacity for Reform', *Public Administration Review*, 6(5) pp. 409-420.
- Bevan, G. and Hood, C. (2006) 'What's measured is what matters: targets and gaming in the english public health care system', *Public Administration*, 84(3) pp. 517-538.
- Bevanda, V., Sinković, G. and Currie, D.M. (2011) 'Implementing a performance measurement system in Croatia', *Measuring Business Excellence*, 15(4) pp. 50-61.

- Bianchi, C. and Rivenbark, W.C. (2012) 'A Comparative Analysis of Performance Management Systems: The Cases of Sicily and North Carolina', *Public Performance and Management Review*, 35(3) pp. 509-526.
- Bititci, U. S., Mendibil, K., Nudurupati, S., Garengo, P., and Turner, T. (2006) 'Dynamics of performance measurement and organisational culture', *International Journal of Operations and Production Management*, 26(12) pp. 1325-50.
- Bourdeaux, C. and Chikoto, G. (2008) 'Legislative Influences on Performance Management Reform', *Public Administration Review*, 68(2) pp. 253-265.
- Bourne, M., Mills, J., Wilcox, M., Neely, A., and Platts, K. (2000) 'Designing, implementing and updating performance measurement systems', *International Journal of Operations and Production Management*, 20(7) pp. 754-771.
- Bourne, M., Neely, A., Mills, J., and Platts, K. (2003a) 'Implementing performance measurement systems: a literature review', *International Journal of Business Performance Management*, 5(1) pp. 1-24.
- Bourne, M., Neely, A., Mills, J., and Platts, K. (2003b) 'Why some performance measurement initiatives fail: lessons from the change management literature', *International Journal of Performance Management*, 5(2) pp. 245-269.
- Bourne, M. (2005) 'Researching performance measurement system implementation: the dynamics of success and failure', *Production Planning and Control*, 16(2) pp. 101-113.
- Boyne, G.A. (2002) 'Public and Private Management: What's the Difference?', *Journal of Management Studies*, 39(1) pp. 97-122.
- Brignall, S. and Modell, S. (2000) 'An institutional perspective on performance measurement and management in the 'new public sector'', *Management Accounting Research*, 11(3) pp. 281-306.
- Brusca, I., Manes Rossi, F. and Aversano, N. (2015) 'Performance Measurement in Italian and Spanish Local Governments: Comparative Policy Analysis', *Journal of Comparative Policy Analysis: Research and Practice*, 10(October) pp. 1-17.
- Bryman, A. and Bell, E. (2007) *Business research methods*. 1st edn. London: Oxford University Press.
- Burnes, B. (1996a) *Managing Change: A Strategic Approach to Organisational Dynamics*. 2nd edn. London: Pitman.
- Burnes, B. (1996b) 'No such thing as ... a "one best way" to manage organizational change', *Management Decision*, 34(10) pp. 11-18.
- Butler, M.J. (2003) 'Managing from the Inside Out: Drawing on 'Receptivity' to Explain Variation in Strategy Implementation', *British Journal of Management*, 14(s1) pp.47-60.

- Carlucci, D., Schiuma, G. and Sole, F. (2015) 'The adoption and implementation of performance measurement process in Italian public organisations: the influence of political, cultural and rational factors', *Production Planning and Control*, 26(5) pp. 363–376.
- Carmona, S. and Grönlund, A. (2003) 'Measures vs actions: the balanced scorecard in Swedish Law Enforcement', *International Journal of Operations and Production Management*, 23(12) pp. 1475-1496.
- Cavalluzzo, K.S. and Ittner, C.D. (2004) 'Implementing Performance Measurement Innovations: Evidence from government', *Accounting, Organizations and Society*, 29(3) pp. 243-267.
- Chang, L.C. (2007) 'The NHS performance assessment framework as a balanced scorecard approach Limitations and implications', *International Journal of Public Sector Management*, 2(2) pp. 101 - 117.
- Cheng, M.I., Dainty, A. and Moore, D. (2007) 'Implementing a new performance management system within a project-based organization: A case study', *International Journal of Productivity and Performance Management*, 56(1) pp. 60-75.
- De Araújo, J. and Branco, J. (2009) 'Implementing Performance-Based Management in the Traditional Bureaucracy of Portugal', *Public Administration*, 87(3) pp. 557-573.
- De Bruijn, H. (2007). *Managing Performance in the Public Sector*. 1st edn. New York: Routledge.
- De Bruijn, H. (2002) 'Performance measurement in the public sector: strategies to cope with the risks of performance measurement', *International Journal of Public Sector Management*, 15(7) pp. 578-594.
- De Waal, A.A. (2003) 'Behavioral factors important for the successful implementation and use of performance management systems', *Management Decision*, 41(8) pp. 688-697.
- De Waal, A.A. and Counet, H. (2009) 'Lessons learned from performance management systems implementations', *International Journal of Productivity and Performance Management*, 58(4) pp. 367-390.
- DeGross, A., Schooley, M., Chapel, T., and Poister, T. H. (2010) 'Challenges and strategies in applying performance measurement to federal public health programs', *Evaluation and Program Planning*, 33(4) pp. 365-372.
- Dibella, A.J. (2007) 'Critical Perceptions of Organisational Change', *Journal of Change Management*, 7(3-4) pp. 231 -242.
- Dopson, S. and Stewart, R. (1990) 'Public and Private Sector Management: the case for a wider debate', *Public Money and Management*, 10(1) pp. 37-40.

- Elg, M. and Kollberg, B. (2009) 'Alternative arguments and directions for studying performance measurement', *Total Quality Management*, 20(4) pp. 409 -421.
- Fernandez, S. and Rainey, H.G. (2006) 'Managing Successful Organizational Change in the Public Sector', *Public Administration Review*, 66(2) pp.168-176.
- Folz, D.H., Abdelrazek, R. and Chung, Y. (2009) 'The Adoption, Use, and Impacts of Performance Measures in Medium-Size Cities Progress toward Performance Management', *Public Performance and Management Review*, 33(1) pp. 63-87.
- Franco-Santos, M. and Bourne, M. (2005) 'An examination of the literature relating to issues affecting how companies manage through measures', *Production Planning and Control*, 16(2) pp. 114-124.
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., and Neely, A. (2007) 'Towards a definition of a business performance measurement system', *International Journal of Operations and Production Management*, 27(8) pp. 784-801.
- Franco-Santos, M., Lucianetti, L. and Bourne, M. (2012) 'Contemporary performance measurement systems: A review of their consequences and a framework for research', *Management Accounting Research*, 23(2) pp. 79- 119.
- Fryer, K., Antony, J. and Ogden, S. (2009) 'Performance management in the public sector', *International Journal of Public Sector Management*, 22(6) pp. 478-498.
- Fábián, A. (2010) 'New Public Management and what comes after', *Current Issues of Business and Law*, 5(1) pp. 41-56.
- Gao, J. (2009) 'Governing by goals and numbers: a case study in the use of performance measurement to build state capacity in china', *public administration and development*, 29(1) pp. 21-31.
- Gao, J. (2015) 'Performance Measurement and Management in the Public Sector: Some Lessons from Research Evidence', *public administration and development*, 35(2) pp. 86-96.
- Gerring, J. (2007) *Case Study Research Principles and Practices*. New York: Cambridge University Press.
- Gianakis, G.A. (2002) 'The promise of public sector performance measurement: Anodyne or Placebo', *public administration quarterly*, Spring pp. 35-64.
- Goh, S.C. (2012) 'Making performance measurement systems more effective in public sector organizations', *Measuring Business Excellence*, 16(1) pp. 31-42.
- Gray, A. and Jenkins, B. (1995) 'From public administration to public management: reassessing a revolution?', *Public administration*, 73(Spring) pp. 75-100.

- Gray, D.E. (2004) *Doing Research in the Real World*. 1st edn. London: Sage Publications.
- Greatbanks, R. and Tapp, D. (2007) 'The impact of balanced scorecards in a public sector environment Empirical evidence from Dunedin City Council, New Zealand', *International Journal of Operations and Production Management*, 27(8) pp. 846-873.
- Greiling, D. (2006) 'Performance measurement: A remedy for increasing the efficiency of public services?', *International Journal of Productivity and Performance Management*, 55(6) pp. 448-465.
- Grindle, M.S. (1997) 'Divergent Cultures? When Public Organizations Perform Well in Developing Countries', *World Development*, 25(4) pp. 481-495.
- Gruening, G. (2001) 'Origin and theoretical basis of New Public Management', *International public management journal*, 4(1) pp. 1-25.
- Hart, C. (1998) *Doing a literature review: Releasing the social science research imagination*. 1st edn. London: Sage Publications.
- Hawke, L. (2012) 'Australian public sector performance management: success or stagnation?', *International Journal of Productivity and Performance Management*, 61(3) pp. 310-328.
- Heinrich, C.J. and Marschke, G. (2010) 'Incentives and Their Dynamics in Public Sector Performance Management Systems', *Journal of Policy Analysis and Management*, 29(1) pp. 183-208.
- Heinrich, C.J. (2002) 'Outcomes-Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness', *Public Administration Review*, 62(6) pp. 712-725.
- Ho, A.T.K. (2005) 'Accounting for the Value of Performance Measurement from the Perspective of Midwestern Mayors', *Journal of Public Administration Research and Theory*, 16(2) pp. 217-237.
- Holmes, J.S., Amin Gutiérrez de Piñeres, S. and Douglas Kiel, L. (2006) 'Reforming Government Agencies Internationally: Is There a Role for the Balanced Scorecard?', *International Journal of Public Administration*, 29(12) pp. 1125-1145.
- Holzer, M. and Kloby, K. (2005) 'Public performance measurement An assessment of the state-of-the-art and models for citizen participation', *International Journal of Productivity and Performance Management*, 54(7) pp. 517-532.
- Hood, C. (1991) 'A public management for all seasons?', *Public Administration*, 69(1) pp. 3-19.
- Hood, C. (1995) 'The 'new public management' in the 1980s: variations on a theme', *Accounting, Organizations and Society*, 20(2-3) pp. 93-109.

Hood, C. (2000) 'Paradoxes of public-sector managerialism, old public management and public service bargains', *International Public Management Journal*, 3(1) pp. 1-22.

IMF (2013) Republic of Yemen Staff Report for the 2013 article Iv Consultation. *IMF*. Available at: <https://www.imf.org/external/pubs/ft/scr/2013/cr13246.pdf>.

Inamdar, N., Kaplan, R.S. and Reynolds, K. (2002) 'Applying the balanced scorecard in healthcare provider organizations', *Journal of Healthcare Management*, 47(3) pp. 179-196.

Jackson, P. (1988) 'The management of performance in the public sector', *Public Money and Management*, 8(4) pp. 11-16.

Jackson, A. (2005) 'Falling from a Great Height: Principles of Good Practice in Performance Measurement and the Perils of top down determination of performance indicators', *Local Government Studies*, 31(1) pp. 21-38.

Jankulović, A. and Škorić, V. (2013) 'Performance management system implementation in a Southeast European transitional country', *Journal for East European Management Studies* pp. 173-190.

Jääskeläinen, A., Laihonen, H., Lönnqvist, A., Palvalin, M., Sillanpää, V., Pekkola, S. and Ukko, J. (2012) 'A contingency approach to performance measurement in service operations', *Measuring Business Excellence*, 16(1) pp. 43-52.

Johansson, T. and Siverbo, S. (2009) 'Explaining the Utilization of Relative Performance Evaluation in Local Government: A Multi-Theoretical Study Using Data from Sweden', *Financial Accountability and Management*, 25(2) pp. 197-224.

Johnsen, Å. (2005) 'What Does 25 Years of Experience Tell Us About the State of Performance Measurement in Public Policy and Management?', *Public Money and Management*, 25(1) pp. 9-17.

Julnes, P.D.L., Holzer, M. and Shick, R. (2009) 'Public Performance Measurement and Management: From Theory to Practice', in *International Conference on Administrative Development: Towards Excellence in Public Sector Performance*. Riyadh: Kingdom of Saudi Arabia Institute of Public Administration.

Julnes, P.D.L. and Holzer, M. (2001) 'Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation', *Public Administration Review*, 61(6) pp. 693-708.

Julnes, P.D.L. and Mixcóatl, G. (2006) 'Governors as Agents of Change. A Comparative Study of Performance Measurement Initiatives in Utah and Campeche', *Public Performance and Management Review*, 29(4) pp. 405-432.

Kaplan, R.S. (2001) 'Strategic Performance Measurement and Management in Nonprofit Organizations', *Nonprofit Management and Leadership*, 11(3) pp. 353-70.

- Kaplan, R.S. and Norton, D.P. (1992) 'The balanced scorecard - measures that drive performance', *Harvard Business Review*, January.
- Kaplan, R.S. and Norton, D.P. (1996). 'Linking the balanced scorecard to strategy', *California management review*, 39(1), pp.53-79.
- Kaplan, R.S. and Norton, D.P. (2001) 'Transforming the balanced scorecard from performance measurement to strategic management: Part 1', *Accounting Horizons*, 15(1) pp. 87-104.
- Kennerley, M. and Neely, A. (2002) 'A framework of the factors affecting the evolution of performance measurement systems', *International Journal of Operations and Production Management*, 22(11) pp. 1222-1245.
- Khan, K. and Shah, A. (2011) 'Understanding performance measurement through the literature', *African Journal of Business Management*, 5(35) pp. 13410-13418.
- Klovieniė, L. and Gimzauskiene, E. (2009) 'Performance measurement system changes according to organization's external and internal environment', *Economics and Management*, 14 pp. 70-77.
- Kravchuk, R.S. and Schack, R.W. (1996) 'Designing Effective Performance Measurement Systems under the Government Performance and Results Act of 1993', *Public Administration Review*, 56(4) pp. 348-358.
- Kroll, A. (2016) 'Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research', *Public Performance and Management Review*, 38(3) pp. 459-486.
- Kuhlmann, S. (2010) 'Performance Measurement in European local governments: a comparative analysis of reform experiences in Great Britain, France, Sweden and Germany', *International Review of Administrative Sciences*, 76(2) pp. 331-345.
- Kwon, M., and Jang, H. S. (2011) 'Motivations Behind Using Performance Measurement: City- wide vs. Selective Users', *Local Government Studies*, 37(6), pp. 601-620.
- Lee, J.W. and Cho, Y.J. (2011) 'In search of a better fit: Exploring conditions for successful performance- based human resources management in U.S. federal agencies', *International Review of Public Administration*, 16(3) pp. 89-112.
- Lee, J.W. and Kim, S.E. (2012) 'Searching For a Strategic Fit an Empirical Analysis of the Conditions for Performance Management Implementation in U.S. Federal Agencies', *Public Performance and Management Review*, 36(1) pp. 31-53.
- Lewin, K. (1951) *Field Theory in Social Science*. New York: Harper and Row.
- Likierman, A. (1993) 'Performance indicators: 20 early lessons from managerial use', *Public Money and Management*, 13(4) pp. 15-22.

- Lilian Chan, Y.C. (2004) 'Performance measurement and adoption of balanced scorecards A survey of municipal governments in the USA and Canada', *The International Journal of Public Sector Management*, 17(3) pp. 204-221.
- Liu, Y.D. (2009) 'Implementing and Evaluating Performance Measurement Initiative in Public Leisure Facilities: An Action Research Project', *Systemic Practice and Action Research*, 22(1) pp. 15-30.
- Mansor, N. N.A., Chakraborty, A. R., Yin, T. K., and Mahitapoglu, Z. (2012) 'Organizational Factors Influencing Performance Management System in Higher Educational Institution of South East Asia', *Procedia - Social and Behavioral Sciences*, 40 pp. 584 - 590.
- Melchor, O. (2008) *Managing Change in OECD Governments: An Introductory Framework. OECD Working Papers on Public Governance*, 12.
- Melkers, J. and Willoughby, K. (1998) 'The State of the States: Performance-Based Budgeting Requirements in 47 out of 50', *Public Administration Review*, 58(1) pp. 66-73.
- Merkens, H. (2004) Selection procedures, sampling, case construction, in '*A companion to qualitative research*', Sage Publications pp. 165-171.
- Micheli, P. and Kennerly, M. (2005) 'Performance measurement frameworks in public and non-profit sectors', *Production Planning and Control*, 16(2) pp. 125-134.
- Micheli, P. and Manzoni, J.F. (2010) 'Strategic Performance Measurement: Benefits, Limitations and Paradoxes', *Long Range Planning*, 43(4) pp. 465-476.
- Mikkelsen, B. (2005) *Methods for development work and research: a new guide for practitioners*. 1st edn. London: Sage Publications.
- Miles, M. and Huberman, A. (1994) *Qualitative data analysis: An expanded sourcebook*. 2nd edn. London: Sage Publications.
- Mimba, N.P., Helden, G.J. and Tillema, S. (2007) 'Public sector performance measurement in developing countries A literature review and research agenda', *Journal of Accounting and Organizational Change*, 3(3) pp. 192-208.
- Mimba, N.P., Helden, G.J. and Tillema, S. (2013) 'The Design and Use of Performance Information in Indonesian Local Governments Under Diverging Stakeholder Pressures', *public administration and development*, 33(1) pp. 15-28.
- Moriani, F., Al-Hammadi, M. and Al-Zawm, A. (2013) 'Options for future form of Government and Decentralization in Yemen Policy Options in times of Change', *UNDP*. Available at:
<http://www.ye.undp.org/content/dam/yemen/DemDov/Docs/UNDP-YEM-Policy Paper.pdf>.

Moynihan, D.P. (2004) 'Why and How Do State Governments Adopt and Implement "Managing for Results" Reforms?', *Journal of Public Administration Research and Theory*, 15(2) pp. 219-43.

Moynihan, D.P. and Pandey, S.K., 2005. Testing How Management Matters in an Era of Government by Performance Management. *Journal of Public Administration Research and Theory*, 15(3) pp. 421-439.

Moynihan, D.P. and Pandey, S.K. (2010) 'The Big Question for Performance Management: Why Do Managers Use Performance Information?', *Journal of Public Administration Research and Theory*, 20(4) pp. 849-866.

Moynihan, D.P., Pandey, S.K. and Wright, B.E. (2011) 'Setting the Table: How Transformational Leadership Fosters Performance Information Use', *Journal of Public Administration Research and Theory*, 22(1) pp. 143-164.

Nath, N. and Sharma, U. (2014) 'Performance Management Systems in the Public Housing Sector: Dissemination to Diffusion', *Australian Accounting Review*, 24(1) pp. 2-20.

Neely, A., Adams, C. and Crowe, P. (2001) 'The performance prism in practice', *Measuring Business Excellence*, 5(2) pp. 6-13.

Neely, A., Gregory, M. and Platts, K. (2005) 'Performance measurement system design A literature review and research agenda', *International Journal of Operations and Production Management*, 25(12) pp. 1228-1263.

Neely, A. (1999) 'The performance measurement revolution: why now and what next?', *International Journal of Operations and Production Management*, 19(2) pp. 205-228.

Newcomer, K.E. (2007) 'Measuring Government Performance', *International Journal of Public Administration*, 30(3) pp. 307-329.

NHS Centre for Reviews and Dissemination (2001), Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews. CRD Report Number 4 (2nd Edition). York.

Northcott, D. and Taulapapa, T.M. (2012) 'Using the balanced scorecard to manage performance in public sector organizations Issues and challenges', *International Journal of Public Sector Management*, 25(3) pp. 166-191.

Nudurupati, S. S., Bititci, U. S., Kumar, V., and Chan, F.T. (2011) 'State of the art literature review on performance measurement', *Computers and Industrial Engineering*, 60(2) pp. 279-290.

OECD (2010) Enhancing Integrity in Public Procurement in Yemen OECD Joint Learning Study. *OECD*. Available at: <https://www.oecd.org/gov/ethics/45916397.pdf>.

Oakland, J.S. and Tanner, S.J. (2007) 'A new framework for managing change', *The TQM Magazine*, 19(6) pp. 572-589.

- Ohemeng, F.L.K. (2009) 'Constraints in the Implementation of Performance Management Systems in Developing Countries the Ghanaian Case', *International Journal of Cross Cultural Management*, 9(1) p. 109–132.
- Ohemeng, F.L.K. (2011) 'Institutionalizing the performance management system in public organizations in Ghana chasing a mirage?', *Public Performance and Management Review*, 34(4) pp. 467-488.
- Ohemeng, F.L.K. (2010) 'The dangers of internationalization and "one-size-fits-all" in public sector management Lessons from performance management policies in Ontario and Ghana', *International Journal of Public Sector Management*, 23(5) pp. 456-478.
- Oliver, P. (2006) Purposive Sampling. In *The Sage Dictionary of Social Research Methods*. Sage pp. 244-245.
- Ostroff, F. (2006) 'Change Management in Government', *Harvard Business Review*, 84(5) pp. 1-9.
- O'leary, Z. (2004) *The Essential Guide to Doing Research*. 1st edn. London: Sage Publications.
- Partington, D. (2002) *Essential skills for management research*. London: Sage Publications.
- Pereira, M.M. and Melão, N.F. (2012) 'The implementation of the balanced scorecard in a school district Lessons learned from an action research study', *International Journal of Productivity and Performance Management*, 61(8) pp. 910-939.
- Perera, S., Schoch, H. and Sabaratnam, S. (2007) 'Adoption of the balanced scorecard in local government organizations: an exploratory study', *Asia-Pacific Management Accounting Journal*, 2(1) pp. 53-70.
- Phusavat, K., Anussornnitisarn, P., Helo, P., and Dwight, R. (2009) 'Performance measurement: roles and challenges', *Industrial Management and Data Systems*, 109(5) pp. 646-664.
- Pidd, M. (2005) 'Perversity in public service performance measurement', *International Journal of Productivity and Performance Management*, 54(5) pp. 482-493.
- Poister, T.H. and Streib, G. (1999) 'Performance Measurement in Municipal Government: Assessing the State of the Practice', *Public Administration Review*, 59(4) pp. 325-335.
- Pollanen, R.M. (2005) 'Performance measurement in municipalities Empirical evidence in Canadian context', *International Journal of Public Sector Management*, 18(1) pp. 4-24.
- Pollitt, C. (2005) 'Performance management in practice: a comparative study of executive agencies', *Journal of Public Administration Research and Theory*, 16(1) pp. 25-44.

Propper, C. and Wilson, D. (2003) 'The use and usefulness of performance measures in the public sector', *Oxford Review of Economic Policy*, 19(2) pp. 250-267.

Radnor, Z. and Lovell, B. (2003) 'Success factors for implementation of the balanced scorecard in a NHS multi-agency setting', *International Journal of Health Care Quality Assurance*, 16(2) pp. 99-108.

Radnor, Z. and McGuire, M. (2004) 'Performance management in the public sector: fact or fiction', *International Journal of Productivity and Performance Management*, 53(3) pp. 245-60.

Rantanen, H., Kulmala, H. I., Lönnqvist, A. and Kujansivu, P. (2007) 'Performance measurement systems in the Finnish public sector', *International Journal of Public Sector Management*, 20(5) pp. 415-433.

Rieley, J.B. and Clarkson, I. (2001) 'The impact of change on performance', *Journal of Change Management*, 2(2) p. 160-172.

Rodgers, M.C. (2011) 'Organizational critical success factors influencing balanced scorecard systems in UK healthcare', *Journal of Management and Marketing in Healthcare*, 4(3) pp. 174-179.

Rousseau, D.M., Manning, J. and Denyer, D. (2008) '11 Evidence in Management and Organizational Science: Assembling the Field's Full Weight of Scientific Knowledge Through Syntheses', *Academy of Management Annals*, 2(1) pp. 475-515.

Sá, P.M. and Kanji, G.K. (2003) 'Finding the path to organizational excellence in Portugese local government: a performance measurement approach', *Total Quality Management*, 14(4) pp. 491-505.

Sanger, M.B. (2008) 'From Measurement to Management: Breaking through the Barriers to State and Local Performance', *Public Administration Review*, 68(S1) pp. 70-85.

Schmidt, C. (2004) The analysis of semi-structured interviews. In *'A companion to qualitative research'*, Sage Publications pp. 253-258.

Šević, Ž. (2005) 'Measuring performance on a local government level in a transitional country: the case of Serbia', *International Journal of Public Sector Management*, 18(7) pp. 582-603.

Sharma, B. and Gadenne, D. (2011) 'Balanced Scorecard Implementation in a Local Government Authority: Issues and Challenges', *Australian Journal of Public Administration*, 70(2) pp. 167-184.

Siddiquee, N.A. (2010) 'Managing for results: lessons from public management reform in Malaysia', *International Journal of Public Sector Management*, 23(1) pp. 38-53.

- Sole, F. and Schiuma, G. (2010) 'Using performance measures in public organisations: challenges of Italian public administrations', *Measuring Business Excellence*, 14(3) pp. 70-84.
- Sole, F. (2009) 'A management model and factors driving performance in public organizations', *Measuring Business Excellence*, 13(4) pp. 3-11.
- Sotirakou, T. and Zeppou, M. (2006) 'Utilizing performance measurement to modernize the Greek public sector', *Management Decision*, 44(9) pp. 1277-1304.
- Stewart, J. and Walsh, K. (1992) 'Change in the management of public services', *Public Administration*, 70(4) pp. 499-518.
- Suddaby, R. (2010) 'Editor's comments: construct clarity in theories of management and organization', *Academy of Management Review*, 35(3) pp. 346-357.
- Taylor, J. (2006) 'Performance Measurement in Australian and Hong Kong Government Departments', *Public Performance and Management Review*, 29(3) pp. 334-357.
- Taylor, J. (2011a) 'Factors Influencing the Use of Performance Information for Decision Making in Australian State Agencies', *Public Administration*, 89(4) pp. 1316-1334.
- Taylor, J. (2011b) 'Strengthening the Link between Performance Measurement and Decision Making', *Public Administration*, 89(3) pp. 860-879.
- Teelken, C. (2008) 'The intricate implementation of performance measurement systems: exploring developments in professional-service organizations in the Dutch non-profit sector', *International Review of Administrative Sciences*, 74(4) pp. 615-635.
- Tillema, S., Mimba, N.P.S. and Helden, G.J. (2010) 'Understanding the changing role of public sector performance measurement in less developed countries', *Public Administration and Development*, 30(3) pp. 203-214.
- Todnem By, R. (2005) 'Organisational Change Management: A Critical Review', *Journal of Change Management*, 5(4) pp. 369 -380.
- Torres, L., Pina, V. and Yetano, A. (2011) 'Performance Measurement in Spanish Local Governments. A Cross-Case Comparison Study', *Public Administration*, 89(3) p. 1081-1109.
- Tranfield, D., Denyer, D. and Smart, P. (2003) 'Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review', *British Journal of Management*, 14(3) pp. 207-222.
- Van Dooren, W. (2005) 'What makes organisations measure? Hypotheses on the causes and conditions for performance measurement', *Financial Accountability and Management*, 21(3) pp. 363-383.

Van Dooren, W. (2011) 'Better Performance Management Some Single- And Double-Loop Strategies', *Public Performance and Management Review*, 34(3) pp. 420-433.

Verbeeten, F.H. (2008) 'Performance management practices in public sector organizations Impact on performance', *Accounting, Auditing and Accountability Journal*, 21(3) pp. 427-454.

Verheijen, T. and Dobrolyubova, Y. (2007) 'Performance management in the Baltic States and Russia: success against the odds?', *International Review Of Administrative Sciences*, 73(2) p. 205–215.

Wallace, M. and Wray, A. (2006) *Critical Reading and Writing for Postgraduates*. 2ed. London: Sage Publications.

Whetten, D.A. (1989) 'What Constitute a Theoretical Contribution?', *Academy of Management Review*, 14(4) pp. 490-495.

World Bank (2008a) Public Expenditure Financial Accountability Assessment: PEFA Assessment. available at:
<http://documents.worldbank.org/curated/en/2008/06/16465532/yemen-public-expenditure-financial-accountability-assessment>.

World Bank (2008b) Yemen Development Policy Review. available at:
<http://documents.worldbank.org/curated/en/2008/04/9436149/yemen-development-policy-review>.

Yang, K. and Hsieh, J.Y. (2007) 'Managerial Effectiveness of Government Performance Measurement: Testing a Middle-Range Model', *Public Administration Review*, 67(5) pp. 861-879.

Yetano, A. (2009) 'Managing Performance at Local Government Level: The Cases of the City of Brisbane and the City of Melbourne', *The Australian Journal of Public Administration*, 68(2) pp. 167-181.

Yin, R.K. (1994) *Case Study Research: Design and Methods*. 2nd edn. Thousand Oaks: Sage Publications.

Yin, R.K. (2002) Applications of Case Study Research', *Applied Social Research Methods Series*, 34.

Zeppou, M. and Sotirakou, T. (2003) 'The "STAIR" model A comprehensive approach for managing and measuring government performance in the post-modern era', *International Journal of Public Sector Management*, 16(4) pp. 320-332.