



Criterion-related Validity Analysis of Performance

Management in Local Government

—A Cross-national Comparison Study

Kristianstad University

The Department of Business Studies

Master Degree Dissertation

Public Administration and International Business

May 2005

Tutors: Christer Ekelund

Annika Fjelkner

Authors: Sun Jun

Xu Xinqiao

Acknowledgments

Having experienced ups and downs, anxiety and hope, perplexity and perspicuity, we come to the goal. In the past five months, we have dedicated our commitment, passion, cooperation, and energy to this dissertation, submitted for master degree of public administration and international business.

This dissertation is the glorious achievement we have ever been dreaming for months long. This dissertation also witnessed our experience of being international students at Kristianstad University in the past months.

This dissertation is dedicated to many people, to whom we are eternally indebted to their great contribution.

To our families, parents and wives, for their examples of love and sacrifice.

To Christer Ekelund, our tutor, for his valuable comments and personal participation, and devotion to our program.

To Annika Fjelkner, our English teacher, for her guiding us through the English language.

To Lisa Nilsson, our program assistant, for her kind and critical help for the connection with Swedish agencies for the survey.

To all those participants.

Kristianstad May, 2005

Sun Jun

Xu Xinqiao

Abstract

Advocated by the New Public Management movement, performance management has been increasingly adopted in local government agencies, aiming to improve government performance and accountability. The purposes of this dissertation are trying to investigate the current practice of performance management and its effectiveness, and find out whether there are differences between Sweden and China, and between different agency levels in terms of performance management and government performance. According to theoretical review, we combine two performance management models into a new one with four stages, and establish a three-dimensional government performance model based on the Balanced Scorecard, with which questionnaires are designed. The results, carried out mainly in China, show that to a large extent the models are consistent with the theory. Nevertheless, a new stage of performance management, named as performance appraisal, emerges combining performance measurement with parts of performance report. It is found that the four stages of performance management are received different attentions in local government agencies. The results also display that two stages of performance management (performance improvement and performance standard-set) have strong causal relations with government performance. Meanwhile, there are significant differences between three agency levels in terms of some performance management stages and one performance dimension, according to the results. As far as the two nations are concerned, Sweden is much better than China at all the four stages of performance management, but only much better at one performance dimension—learning and growth—than China.

Keywords: performance management, performance, government agency, stakeholder

Table of Content

Chapter 1	Introduction	1
1.1	Background.....	1
1.2	Introduction of performance management.....	2
1.3	Problem	2
1.4	Purpose.....	2
1.5	Limitations	3
1.6	Research Questions	3
1.7	Outline.....	4
Chapter 2	Methodology	5
2.1	Choice of methodology.....	5
2.2	Collections of data	6
2.2.1	Secondary data	6
2.2.2	Primary data.....	7
2.3	Scientific approaches	7
Chapter 3	Theoretical Framework.....	9
3.1	New Public Management Movement.....	9
3.1.1	Background and development of New Public Management movement	9
3.1.2	The components of the New Public Management	10
3.1.3	The theoretical foundation of the New Public Management movement ..	11
3.1.4	Empirical studies on the validity of the NPM.....	13
3.2	Performance management reform.....	14
3.3	Performance management.....	15
3.3.1	Definition and procedures of performance management	15
3.3.2	How to make performance management effective	17
3.3.3	Validity analysis of performance management.....	18
3.3.4	Summary.....	18
3.4	Important procedures of the performance management.....	19
3.4.1	Set performance standards.....	19
3.4.2	Select performance indicators.....	20
3.4.3	Performance measurement.....	21
3.5	Performance management model	23
3.6	Organizational performance	25
3.6.1	Terminology of the performance.....	25
3.6.2	Private sector performance	26
3.7	Government performance	27
3.7.1	Government performance structure.....	27
3.7.2	Balanced scorecard	29
3.7.3	Stakeholder theory	30
3.7.4	Monetary versus Non-monetary indicators	31
3.7.5	Summary.....	32
3.8	The development of our model.....	33
3.8.1	The performance management variable	34

3.8.2	The government performance scale.....	39
3.8.3	Research hypotheses	42
Chapter 4	Empirical Method	45
4.1	Research Strategy.....	45
4.2	Sample	47
4.3	Limitations.....	48
4.4	The questionnaire.....	48
4.5	Response rate.....	49
4.6	Reliability.....	50
4.7	Validity.....	50
4.8	Validity generalization	51
Chapter 5	Analysis of the Survey.....	52
5.1	Introduction of the statistical methods.....	52
5.2	Components of the respondents	53
5.3	Factor analysis results.....	56
5.3.1	Factor analysis results of the performance management scale	57
5.3.2	Factor analysis of government performance scale	62
5.4	Descriptive analysis.....	66
5.4.1	General descriptive analysis.....	66
5.4.2	Differences of the ratings between job ranks.....	67
5.5	Correlation and Regression analysis	68
5.5.1	Correlations between independent and dependent variables.....	68
5.5.2	Regression	69
5.5.3	Summary.....	70
5.6	Differences on the ratings between agency levels	71
5.6.1	Descriptive analysis.....	71
5.6.2	Differences between agency levels.....	72
5.6.3	Summary.....	75
5.7	Comparison between two nations	76
5.7.1	Mean comparison on the four factors of the performance management	76
5.7.2	Mean comparison of the three factors of the government performance	77
5.8	Summary of the whole analysis.....	78
Chapter 6	Conclusions	80
6.1	Summary of the dissertation	80
6.2	Applicability of the models.....	82
6.3	Methodological criticism	84
6.4	Future research	84
6.5	Managerial implications to the Chinese local government.....	85
References.....		88
Appendices.....		95
Appendix 1.....		95
Appendix 2.....		99
Appendix 3.....		104
Appendix 4.....		106

Chapter 1 Introduction

In the first chapter the background of the dissertation is introduced. The research problems and the purposes are discussed. Further, the limitations and the research questions are described. In the end, the outline is presented.

1.1 Background

In recent years, China has deepened and widened governmental reforms in order to meet an increasing public demand in terms of efficiency and accountability. Meanwhile, performance management, one critical managerial tool in the public sectors prevailing in western countries for two decades, has received great attention in the last years in Chinese local governments. As, public servants in a local government, we are very interested in this topic and wonder what the real meaning of the performance management, how to implement the performance management, and what the effects of performance management, are. Our master degree program of Public Administration and International Business provide us with an opportunity to do a research in this field. After reading many relevant articles on the subject, we are gradually becoming familiar with performance management, from theory to practice. Furthermore, we notice that most of scholars share a similar idea about performance management being integrated with a set of procedures. But, being an integrated managerial tool, it has not yet been implemented systematically in many government sectors. In addition, confirmation of the criterion-related validity of performance management is still in question with empirical data. Meanwhile, very little attention has been paid to the comparative analysis at different government levels. On the other hand, few cross-national comparison researches have ever been implemented. Hence, we finally decided to focus on the validity of performance management and on all international comparison as well.

1.2 Introduction of performance management

Performance management, being a performance reform, has become major roles of the operation in the government organizations throughout the western countries aiming to improve organizational performance and accountability (Metzenbaum, 1998). All the OECD member countries have taken performance management reform so far. In the late 1980s, Australia and New Zealand became the first two countries adopting performance management. In the first half of the 1990s, another seven countries followed the trend including Sweden and Austria. Germany and Switzerland follow on the heels of their counterparts later (OECD, Governing for performance, 2004).

Generally, performance management is comprised of four management stages, including performance standard setting, performance measuring, performance reporting, and performance improving, according to most scholars.

1.3 Problem

Although performance management has been implemented, partly or wholly, in local or central governments all around the world in recent years, the academic area has not yet made sufficient research with respect to the concept, the process and the effects of performance management. The problem is that there is still not enough empirical evidence to testify its structure and validity. Moreover, international comparison analysis has not received enough attention in performance management area.

1.4 Purpose

The purpose of our research is to investigate the current practice of implementation of performance management in local government, both in Sweden and China. Further, we want to see whether there is casual relationship between performance management and government performance. In other words, what is the criterion-related validity of performance management? Finally, we attempt to discover the similarities and

differences among these practices through a cross-national comparison.

1.5 Limitations

Many researchers have studied performance management and government management. However, due to the limited time, we can only focus on some of them. It becomes our first limitation. Secondly, considering the possibility of our research, we only make a survey among public servants in three municipalities in Skåne, Sweden and Ningbo municipality in China. So, these samples do not suffice to represent the two countries adequately. The third limitation is that the dimensions of government performance employed in this research do not include financial indicators. Hence, the study cannot show the whole picture of government performance. The fourth limitation is the way of selecting the respondents, especially the Chinese respondents. We just asked our friends to motivate their colleagues to participate in our survey. Therefore the distribution of the respondents is confined to a comparatively small area.

1.6 Research Questions

The dissertation is based on the following Research Questions:

- Is there any high criterion-related validity of performance management, or in other words, any significant causal relationship between performance management and government performance?
- What are the current practices the extent to which the local government agencies has been done to manage performance in terms of the four stages of performance management?
- Are there any differences in the implementation of performance management among three levels in the Chinese local government agency?

- Are there any differences in the implementation of performance management and the rating of government performance between Sweden and China?

1.7 Outline

The dissertation has the following outline.

In chapter 2, we will present the methodological strategy used in our dissertation, the method of data collection data and the scientific approach.

In chapter 3, the theoretical framework will be introduced. Firstly, we simply review the New Public Management movement and analyze its implication on performance management in government. Secondly, definitions, contents, and models of performance management and government performance are discussed. Finally, based on a theoretical review, we integrate Mwita's and Landrum and Baker's model into a new model of performance management, and select balanced scorecard as the model of government performance. We present our hypotheses in the end of the chapter.

In chapter 4, we present our empirical method. The research strategy, the samples and limitations of the survey are discussed. Further, the questionnaire and response rate is presented. The chapter is ended with validity, reliability, and generalization analysis.

In chapter 5, we analyze the results of the survey and evaluate the hypotheses.

In chapter 6, we present the conclusion of our research. The dissertation is summarized briefly. Suggestions for future research are also presented.

Chapter 2 Methodology

In this chapter, we discuss the methodology that is used in our research. We continue with presenting the collection of the secondary and primary data. In the end, the scientific approach of the dissertation is introduced.

2.1 Choice of methodology

Generally, researchers use a deductive or inductive approach to fulfill their research. As to the difference of these two research approaches, Saunders et al. (2003) explains that "...your research should use the deductive approach, in which you develop a theory and hypothesis (or hypotheses) and design a research strategy to test the hypothesis, or the inductive approach, in which you would collect data and develop theory as a result of your data analysis." It means that a deductive approach is a way where researchers develop hypotheses based on an existing theory and tests them with quantitative data, aiming to verify causal relationships between variables and to generalize as well. On the other hand, an inductive approach is a method that researchers use to collect qualitative data to develop a new theory when/where there is no established theory. Furthermore, in the inductive approach, researchers have "a close understanding of the research context" and "less concern with the need to generalize" (Saunders et al, 2003: 86-89).

Which approach is selected depends on the existence of relevant theories, the time available for researchers and the extent of the risk that researchers can bear (Saunders et al, 2003: 90). The main purpose of our dissertation is to analyze the criterion-related validity of performance management. The first step is to review the existing literature with respect to performance management (independent variable) and government performance (dependent variable) carefully and closely. In the second step, a research model combining independent and dependent variables is put forth based on two existing models. Subsequently, research hypotheses are outlined according to the research model. Meanwhile, a questionnaire survey is chosen as a

fundamental method to collect information. Finally, the results of the survey are evaluated to testify the hypotheses. Taking the research steps into account, a deductive way is prevalent in our research.

2.2 Collections of data

2.2.1 Secondary data

The main purpose of our research is to find out whether any causal relationship exists between performance management and government performance. To fulfill the research, we have collected articles and books from the library and the Internet as secondary data. A research model was established based on literature review.

Firstly, we make a brief introduction of the New Public Management, which strongly advocates the implementation of the performance management in the public sectors.

Secondly, definitions of the performance management put forth by different scholars are reviewed (OECD 1995; Jackson and Broom et al. 1999; Roger 1994). Furthermore, we make a brief discussion about the procedures of performance management (Mwita 2000; Bennett and Hill 2001; Whittaker 2003; Bevan and Thompson 1991; Amaratunga and Baldry 2002). The most important procedures, such as setting performance standards, selecting performance indicators, and performance measure are described in detail (TPPO 2003; Kates et al 2000; Leuven 1997; Kravchuk and Schack 1996). In the end, two performance models are singled out to be a basis for the establishment of our own research model (Mwita 2000; Landrum and Baker 2004; Turning Point National Program Office 2002).

Finally, we review the relevant literature with respect to government performance. After describing the definition and structure of government performance, the two most famous models are discussed in depth. The first is Fitzgerald six-dimension model (Fitzgerald et al 1991) and the second is the balanced scorecard (Kaplan and Norton 1992). Considering the internal structure and the feasibility of the data

collection, the balanced scorecard is selected to be the framework of the dependent variable.

2.2.2 Primary data

Firstly, the research model is operationalized into forty standardized questions. Then questionnaire surveys both online and with paper-pen method are elaborately designed to collect data both from Sweden and China. The Swedish sample is mainly concentrated on three municipalities in Skåne. Correspondingly, public servants from government agencies at three levels are regarded as the sample of Chinese database. E-mail with Swedish explanations is sent to personnel of the agencies in the three Swedish municipalities. Meanwhile, we call or send e-mails to our friends working as government officials, asking them and their colleagues to participate in the survey. As a result, 196 Chinese and 10 Swedish respondents have taken part in the survey, and their answers can constitute the primary data for our dissertation.

2.3 Scientific approaches

There are three kinds of research philosophies: positivism, interpretivism and realism. The positivistic philosophy refers to the researcher who holds an objective attitude to interpret collected data in a value-free manner. The researcher also tries to stay neutral and affect or to be affected by what is being studied as little as possible. Furthermore, the outcome of this kind of research is pursued to be replicable (Saunders et al., 2003: 83).

The interpretivistic philosophy is the one that focuses on analyzing the details of events in order to understand the reality or perhaps a reality working behind them. Researchers advocate that this philosophy is opposite to positivism and maintain that the social world cannot be generalized because of its uniqueness and complication (Saunders et al., 2003: 84).

The realistic philosophy is a philosophy that argues “a reality exists that is

independent of human thoughts and beliefs.” Under this philosophy, the researcher believes that large-scale social forces can affect people’s perception (Saunders et al., 2003: 84).

As for our research, we think it is based on the positivistic philosophy because we try to conduct our research with an objective and neutral attitude.

A research can also be qualitative or quantitative according to the different kinds of the research data. A qualitative research means the data the researcher collects is not standardized, and is expressed with words. While a quantitative research means that the research data collected by standardized questions and reflected by a series of numbers (Saunders et al. 2003). In our research, the data is collected by questionnaire and analyzed with statistics. Therefore, our research is quantitative. In addition, the subjects of our research are particular phenomena at a particular time, making our research a cross-sectional study.

Chapter 3 Theoretical Framework

The theoretical framework is presented. First, we introduce the New Public Management theory and movements. Second, we discuss the performance management, focusing on its key processes. Different definitions and models are presented. Third, government performance is discussed. Fourth, we modify the existing models and develop our model for the research.

3.1 New Public Management Movement

3.1.1 Background and development of New Public Management movement

Since the late 1970s, the developed countries are facing increasing fiscal pressure and public expectation on the accountability and efficiency of the government. To cope with the crisis, these countries have initiated a range of government reforms that differ from the traditional public administration.

After having studied these government reforms throughout the world, both developed and developing countries, scholars point out that the reform movement indicates that traditional bureaucratic administration had been replaced by a new management paradigm based on market scheme. This new paradigm aims at economy, efficiency and effectiveness through economization within the public sectors. Scholars label it with different terms, such as New Public Management (NPM), Market-based Public Administration, and Post-bureaucratic Paradigm and so on (Yan, 2003). Among all these terms, NPM is widely accepted (Dunsire, 1995). König (1997) argues that the NPM is a “popularized mixture of management theories, business motivation psychology and neo-liberal economy.”

Metcalf (1998) thinks the NPM is an umbrella term, with a wide range of meanings,

such as organizational and management design, the application of new institutional economics to public management, and a pattern of policy choices (Barzelay, 2002).

As for the driven forces of the NPM, Lin (1999) outlines a list combined of economical, societal, political and technological factors. The first force is the increasing dissatisfaction with the public sectors in terms of efficiency and quality. The second force is the introduction of economics theories into the operation of the public sectors. The third force is the impact of the management methods formerly used in the private sector on the public one. The fourth force is the development of technology, especially the information technology, which makes it possible for the public sectors to introduce different kinds of managerial reforms.

3.1.2 The components of the New Public Management

After having studied the reform experiences that different countries go through, scholars put forth different components of the NPM.

Larbi (1999) summarizes that the common characteristics of the NPM are the introduction of market and competition in the public services, an increased use of outputs control, the introduction of performance management, an emphasis on customer orientation and a decentralization of management within the public organizations.

Hood (1991) argues that the NPM is an international phenomenon concerning government reform. He describes seven components of the NPM as follows: “(1) an stress on hands-on professional management skills within the public sector; (2) a emphasis on private-sector-style management practice, such as mission statements, and the development of corporate plans; (3) clear standards and measures of performance by setting up goal, defining target and selecting measurable indicators; (4) a stress on the control of outputs, allocating recourses according to measured performance; (5) a shift to greater competition and mixed provision so as to low costs;

(6) an introduction of decentralization or desegregation of units into quasi-contract; (7) a stress on greater parsimony in recourses use, efficiency, cost-cutting”(Hood, 1991).

Osborne and Gaebler (1992) point out that the NPM could reinvent the government towards: “(1) catalytic government, (2) decentralized government, (3) competition government, (4) community government, (5) mission-driven government, (6) customer-driven government, (7) result-oriented government, (8) enterprising government, (9) market-oriented government, (10) anticipatory government”.

OECD (1991) continuously tracks various government reforms in its member countries in the 1990s, and concludes that “a new paradigm for public management”, namely the New Public Management has emerged (Kickert, 1997). According to the OECD (1991), the NPM has eight elements: “(1) devoluting authority and providing flexibility; (2) ensuring performance, control, accountability; (3) developing competition and choice; (4) providing responsive service; (5) improving the management of human resources; (6) optimizing information technology; (7) improving the quality of regulation; (8) strengthening steering functions at the center”.

From what has been mentioned above, we conclude that all the researchers have found that a performance related managerial method is a key element of the NPM initiatives to improve the performance and accountability of the public sector.

3.1.3 The theoretical foundation of the New Public Management movement

Mueller (1989) notes that two main theoretical sources underpin the theoretical foundation of the NPM. Hood (1991) argues that the NPM developed from “a marriage of two different streams of ideas.” One is the new institutional economics theory, which is built on public choice theory, agent theory and transaction-cost theory. Another is “managerialism”, which refers to the management ways used in business

administration (Aucoin, 1990; Hume, 1981; Merkle, 1980; Pollitt, 1993).

The Public Choice Theory. The core assumption of this theory is that the human behavior is motivated by self-interest (Buchanan, 1987; Boston, Martin, Pallot and Walsh, 1996). The theory argues that governmental officials have their own specific goals; hence they focus on maximizing the budget of their agency. Meanwhile, politicians aim to gain more votes (Yan, 2003). Consequently, the government agencies increasingly demand more resources and economic depression will ensue. The theory also notes that one of the fundamental features of the government agency is the non-market characteristic of its output, which is hard to measure (Yang, 2002). It is also argued that the monopoly of the public goods production by a bureaucratic system results in a lack of competition and supervision. The lack of competition imposes no pressure of cost reduction and efficiency improvement on the public sectors. The lack of supervision means the producer can manipulate the supervisor (e.g. the public, the superior agent) because of its dominance in the production process (Chen, 2003).

The Agent Theory. The theory argues that the essence of the social and political life is the contracts between the principals and agents (Liu, 2002). However, there are many problems deriving from different and divergent interests or goals between the two parties (Boston, Martin, Pallot and Walsh, 1996: 18-19). Meanwhile, the agent always has “an informational advantage” over the principal (Ferris and Graddy, 1998). On the contrary, the principal is not accessible to important information, which makes it very difficult to monitor the agent (Moe, 1984: 765-766). Others argue the most important thing between the two parties is how to avoid the goal conflicts (Perrow, 1986: 13-14; Waterman and Meier, 1998: 183-185). Ferris and Graddy (1998) point out that the principal has to take concrete steps to guard against the agent’s opportunistic behavior e.g. effectively monitor and enforce contract.

The Transaction Cost Theory. The transaction costs are the costs related with transactions between parties. For instance, in a hierarchical organization, i.e. a

government agency, the transaction costs relate with “transmitting information within the hierarchy, with implementation and with communication”, determined by the “frequency of the transaction, its complexity, the extent of uncertainty, the extent to which performance is measurable (Ferris and Graddy, 1998). They argue that information asymmetry is always inevitable during the transaction, which allows “opportunistic behavior and its associated costs”. The theory advocates introducing the market scheme and managerial methods used in the private sectors into the public sectors to reduce the transaction costs.

The Managerialism theory. Pollitt (1993) defines the managerialism as a set of faith and practice or a kind of ideology. He argues that effective management could overcome a broad of economic and social symptoms. Managerialism theory maintains that different organizations, private or public, have similar managerial problems and demands. Farnham and Horton (1996) argue that the public sectors should introduce competitive scheme, which demands the public employees to pay much more attention to cost, quality, and effectiveness. They point out that a market-oriented organization is much more efficient than a traditional bureaucratic one.

3.1.4 Empirical studies on the validity of the NPM

Many researchers have carried out a number of studies to investigate the relationship between the NPM reforms and government performance. Boyne et al. (2003) select independent variables from the reforms as a basis. On the other hand, the dependent variables are also specified for assessing the consequences of the ‘deliberate’ reform. Boyne et al. (2003) single out three criteria—efficiency, responsiveness, and equity—and name them as “public service improvement”. Pollitt lists broader criteria, such as “savings, improved processes, improved efficiency, greater effectiveness, and an increase in the overall capacity/flexibility/resilience of the administrative system as a whole”(Pollitt, 2000). Boyne (2003) analyzes which factor could contribute to service improvements. He (2003a) also makes a meta-analysis and ensures the causal

structure between the reform and the government performance.

However, there still has very little empirical evidence assessing the contributions of the reform initiated by the NPM to governmental performance. Pollitt (2000) notes that there is still short of rigorous empirical verification of the causal relationships between the NPM and government performance. Boyne et al. even assert that “the academic community has not taken seriously to evaluate the public management reforms”(Boyne et al 2003).

On the other hand, many studies in line with the NPM theory are carried out in one specific nation, such as USA, Australia, or undertaken independently in several different countries lacking of direct comparison. In recent years, public management variables are increasingly introduced in international studies to explain performance (Fobes and Lynn, Jr., 2004). Now researchers begin to enlarge their academic views and focus on the new public management issues and the government performance; cross-country or international studies prevail (Pollitt, 2000, Pollitt and Bouckaert, 2000, Boyne, 2003, Boyne *et al.*, 2003).

3.2 Performance management reform

In recent years, the worldwide movement of the NPM assumes that government agency can (and should) follow managerial techniques from the private sector, which uses measurement and evaluation of the organizational performance to reach its goals (LOGOTRI, 2001). Though NPM differentiates in terms of operational methods across countries, the convergence is that one of its focuses is performance appraisal (Bevir et al., 2003).

To meet with the higher public expectation of accountability and demands for efficiency and effectiveness in the operations, the governmental organizations begin to focus on and implement performance management reform (Hood, 1995) with an explicit purpose to “increase transparency and improving government operations”

(Martin and Singh, 2004).

The ever-increasing focus on the government performance management reform takes place at all levels, central and local government, to meet “public and elected officials’ demands for accountability” (Berstein, 2000). Kloot and Martin (2000) argue that most of the attentions on performance management reform are in local governments. Generally, this reform takes place at country, agency, and project levels (Binnendijk, 2001).

Some commentators argue that this reform has already developed since the late 1970s in the government sectors concerning with “collecting, reporting, and appraising organizational performance”; they name this reform as an “industry” (Holloway, 1999; Rouse, 1993, 1999).

In the UK, the symbol of the reform is the Financial Management Initiative, which was introduced in the early 1980s. In USA, the reform was culminated in the American Government and Performance Results Act of 1993.

3.3 Performance management

3.3.1 Definition and procedures of performance management

3.3.1.1 Definition

The OECD defines performance management as a system in which “program performance objectives and targets are determined, managers have flexibility to achieve them, actual performance is measured and reported, and this information feeds into decisions about program funding, design, operations and rewards or penalties” (OECD, Governance in Transition, 1995).

Many researchers give different definitions in words but they are all similar in meaning definitions. Jackson and Broom et al. (1999) define performance management as a method of evaluating the “progress of a public program or activity

in achieving the results or outcomes that stakeholders expect”. Rogers (1994) notes it is an “integrated set of planning and review procedures”, linking individual performance with the strategy of the organization. Edis (1995) maintains that any integrated, systematic approach aiming to improve organizational performance can be called performance management.

3.3.1.2 Procedures

Being a system, the performance management has a set of steps, or procedures.

Mwita (2000) notes that performance management is an integrated set of “planning and review procedures”. Bennett and Hill (2001) point out that performance management systematically combines “strategic planning, performance measurement, program evaluation, and performance budgeting processes”. Whittaker (2003) notes that performance management consists of strategic plans, articulating goals and objectives, and execution management. Boland and Fowler (2000) argue that it is composed of “performance measures, performance indicators, performance appraisal and review, value for money and, more recently, quality assurance”. Others think (Bouckaert, 1993, Wholey and Hatry, 1992, Kravchuk and Schack, 1996, and Wholey, 1999) that performance management is combination of “administration of performance measurement, selection of performance measures, timing of performance measurement, and performance benchmarking”.

According to Bevan and Thompson (1991), performance management is composed of five elements: “(1) A shared vision of organizational objectives, or a mission statement, communicated to all employees; (2) Individual Performance Management targets related both to operating unit and wider organizational objectives; (3) Regular formal reviews of progress towards these targets; (4) The review process used to identify employee training, development and reward outcomes; (5) Evaluation of effectiveness of the whole process and its contribution to overall organizational performance to allow changes and improvements to be made”.

Amaratunga and Baldry (2002) report that a performance management system has several steps, including (1) Active communication within and outside the organization; (2) Develop organizational strategic goals; (3) Demonstrate improvement needs within the organization; (4) Facilitate learning and improvement; (5) Co-ordinate responsibilities with employees; (6) Attempt to implement the best practices; (7) Training improvement techniques; (8) Establish a reward system to foster improvements; (9) Integrate measurement into the organization; and (10) Change organizational culture.

3.3.2 How to make performance management effective

Binnendijk (2001) points out two main goals of the performance management system: accountability (e.g. holding managers accountable for the performance) and management improvement (e.g. learning and decision- making).

Mercer (2004) maintains that an effective performance management system can guarantee an “agency’s administrative and support functions” in a direct and explicit way to meet the agency’s strategic goals; and it also could motivate and enable managers to maximize performance at minimize cost.

Tat-Kei Ho and Coates (2002) argue that citizen should and must participate in the performance management system to ensure the accountability so that the public programs could address the citizen’s concerns. In other words, citizens, elected officials, together with public staffs should cooperate with each other.

Brown et al. (1996) provide some suggestions for the public sectors to establish a successful performance management system: “(1) Leadership commitment; (2) A desire for accountability; (3) A conceptual framework; (4) Strategic alignment; (5) Knowledgeable and trained staff members; (6) Effective internal and external communication; (7) A positive not punitive culture; (8) Rewards linked to performance; (9) Effective data processing systems; (10) A commitment to and plan for using performance information”.

However, Propper and Wilson (2003) point out that there are some issues still in question with respect to the usefulness of the performance management. First, the purpose of the performance management may differ with government levels in which it is undertaken. Second, multiple stakeholders mean that “a single performance management may be not sufficient”. Third, the publication of performance management to the public may depend on the environment. Fourth, performance management is consists of short-term measures, which may increase the possibility that performance management will misdirect activity while ignoring the long-term benefits.

3.3.3 Validity analysis of performance management

Though performance management has been widely used in government sectors, there is still a lack of evidence to prove its usefulness in the academic area (Propper and Wilson, 2003). In the political field, politicians may have no interests in spending money to undertake evaluations of the impacts of performance management, which will only be revealed years later. Propper and Wilson (2003) argue that non-experimentation of government policy make it difficult to assess the criterion-related validity of performance management. Moreover, some studies even negatively show the effectiveness of the performance management (Schellhardt, 1996).

3.3.4 Summary

In this study, performance management is an independent variable, or the predictor of government performance. Hence, the construct is of vital importance. There are so many different definitions and models, trying to make a detailed interpretation of the performance management. After studying these models closely, we find that all the definitions and models seem to show the same thing. Based on these findings mentioned above, we draw the conclusion that performance management is a set of processes or a procedure, comprising of setting standards, selecting indicators,

measuring, reporting, and improving based on the report. The ways or techniques the performance management system adopts are not the focus of this study.

3.4 Important procedures of the performance management

According to the researches mentioned above, setting performance standards, selecting performance indicators and measuring performance are the key processes of the performance management system.

3.4.1 Set performance standards

Performance standard is a kind of statements added to performance indicators to specify the level of acceptability for each indicator (Handbook of State of Oklahoma, 1998). TPPO (2003) thinks that it is “objective standards or guidelines that are used to assess an organization’s performance”.

To set a performance standard is one of the most important steps in the performance management system (Courty and Marschke, 2004). TPPO (2003) thinks it may be set based on “national, state, or scientific guidelines; by benchmarking against similar organizations; based on the public’s or leaders’ expectations; or other methods.” Courty and Marschke (2004) further outline four approaches to set performance standards, including estimation of production function, past performance, relative performance evaluation, and negotiation of standard.

Keehley et al. (1997) regard benchmarking as an important method in setting the performance standard. Fischer (1994) notes that the benchmarking concerns with comparing performance with others’ so as to obtain new ways and ideas. Cook (1995) makes a similar definition and considers the benchmarking aims at improving the organizational performance through identifying, understanding, and adopting best practices from similar organizations or industries.

LOGOTRI (2001) concludes that governments at different levels may take a little bit

different way of benchmarking. Local Authorities mainly focus on selecting benchmarking “partners”, other organizations that adopt the same performance indicators, to make a comparison across organizations. On the other hand, central government agencies pay more interests on selecting “best practice” to formulate standards for the local governments to achieve.

3.4.2 Select performance indicators

Performance indicators are “specific values of each measure that describe position within that particular dimension of performance” (Kates et al., 2000), or it is a measure that “summarizes relevant to a particular phenomenon or a reasonable proxy for such a measure” (McQueen and Noak, 1988). Kates et al. (2000) regard it as “specific values of each measure that describe position within that particular dimension of performance”. Others define that it summarizes the focus of performance goals and measures, often used for communication purposes and preceding the development of specific measures (Turning Point, 2002).

3.4.2.1 Function of the performance indicators

Burningham (1992) maintains that the performance indicator has three main functions, such as showing accountability, controlling, and development. Dunn (1994) thinks that performance indicators have three functions: (1) simplify the complexity, (2) quantify development, and (3) fulfill communication. Similarly, Rouse (2003) agrees that effective performance indicators are fundamental to the success of the performance management, through improving internal management within the public sector, improving the accountability, and guiding the employees to a right direction.

3.4.2.2 Types of performance indicators

Performance indicators can be divided into two categories: quantitative indicator and behavioral indicator. The behavior indicator means the extent to which the organization or its employees behave in a specific way (Liu, 2003).

Managers always appreciate the quantitative indicators. However, organizations always obtain the performance through evaluating the extent of people's behavior in reality (Liu, 2004).

3.4.2.3 Select the performance indicator

Zhu (2003) argues that the core concept of performance management is to design the performance indicators.

Dunn (1981) provides seven principles for selecting performance indicators, which include relevance, significance, validity, reliability, objectivity, timeliness, and usability.

Carlin (2001) maintains that performance indicators should capture six key factors, including correlative, controllable, comprehensible, contemporaneous, consistent, and constrained.

Thiel and Leeuw (2002) argue that the performance management system should encompass enough numbers of indicators to ensure the credibility. They suggest that external parties rather than organizations themselves should develop the indicators, and indicators should take political or democratic outcomes into consideration.

There are two main methods to select performance indicators, benchmarking and strategic planning (Liu, 2004). The strategic planning is a systematic procedure that integrates organizational goal, strategy, resources, and administrative process so that the organization can continuously monitor and manage these important dimensions (Denning, 1971, Huang, 2000).

3.4.3 Performance measurement

3.4.3.1 Definition and contents of the performance measurement

A vast array of literature claim that these two terms seem to be “interchangeable”, or that performance measurement is one part of performance management (LOGOTRI,

2001). Some scholars share with the same idea. Leuven (1997) and Binnendijk (2001) maintain that performance measurement is the “logical prerequisite” and dominant part or key component of the performance management system.

Tønneisen (2004) regards the performance measurement as a process of “collecting and tracking the data used in performance management”. Kloot and Martin (2000) consider that performance measurement mainly focuses on the measurement process. On the other hand, Scotti (2004) argues that the fundamental part of the performance measurement is to establish clear accountabilities and expectations, and “output” reports.

Leuven (1997) outlines a list of instruments of performance measurement, which encompassing “performance indicator systems, accrual cost accounting concepts, performance budgeting, performance monitoring/controlling, benchmarking procedures, quality assessments and awards, general performance awards and competitions”.

3.4.3.2 Select measurement tool

Considering the complexity of public goods or services, the government agencies face great difficulty in selecting performance measurement tools. Generally, the measurement tools can be divided into five categories: output measures, efficiency measures, outcome measures, benchmarking review, and cost-effectiveness measures (Liu, 2003).

3.4.3.3 Effectiveness of performance measurement

The NPM literature advocates using performance measurement so as to show and gain public accountability (Epstein, 1988; GASB, 1994; Ammons, 1995; Holzer and Halachmi, 1996). Kaboolian (1998) points out that performance measurement also serves a useful tool to evaluate the effectiveness of the NPM initiatives, e.g. emphasizing competition, empowerment of managers.

Ittner and Larcker (1998) mention that in recent years the government organizations have paid considerable attention to performance measurement with an ambitious purpose to “increase accountability and improve decision-making”. Benowitz and Schein (1996) believe that it has been widely believed to the government to gain public confidence through the means of performance measurement.

Some think the introduction of performance measurement is a mean to demonstrate accountability (Feller, 2002) and rebuild public perceptions of government in a positive direction (Berstein, 2000).

Others regard it as a way to improve performance. Holzer and Yang (2004) note that performance measurement is a theoretical ground for sophisticated performance improvement strategy in the government sectors. Lee (1997) regards it as a useful managerial tool to “monitor, improve and update management practices”. Hoogenboezem (2004) notes that the introduction of the performance measurement makes it possible for the government to assess the performance and to “take punitive action” toward those who fail.

Nonetheless, there is still very limited effort and empirical data assessing the relationship between “contributions of public management and government performance” irrespective of the increasing concern of public management reform by governments throughout the world (Forbes, 2004). The available empirical data show quite mixed results of the role of the performance measures (Christensen and Yoshimi, 2003). The NPM literature itself also lacks of empirical data to prove the effect of this performance reform (Berstein, 2000).

3.5 Performance management model

The performance management model is defined as a “systematic and data-oriented” approach to improve the organizational performance (Daniels, 1989).

Mwita (2000) puts forth a five-factor performance management model, consisting of

five dimensions: mission statement, strategies and plans, action planning, performance recognition, and management accounting information system.

Table3.1 Mwitwa’s Performance Management Model

Mission Statement	<ul style="list-style-type: none"> – Purpose of existence – Community expectations – Values
Strategies and plans	<ul style="list-style-type: none"> – Corporate objectives – SWORT analysis – Training, development and retention policies
Action Planning	<ul style="list-style-type: none"> – Operational objectives – Key tasks & responsibilities – Resource allocation – mapping – Performance target setting – Agreement & performance contracting
Performance recognition	<ul style="list-style-type: none"> – Measurement of performance – Rewards and sanctions – Training needs assessment
Management Accounting Information System	<ul style="list-style-type: none"> – Feedback/feed forward loops – Performance information-set – Timely availability for planning, decision making and control – Suggestions box

Landrum and Baker (2004) put forward a four-essential-component model for performance management, consisting of performance standards, performance measures, progress reports, and quality improvement.

Performance standards mean the “establishment of performance standards, targets, and goals and relevant indicators”. Performance measures refer to the “application and use of performance indicators and measures”. Reporting of progress is the “documentation and reporting of progress in meeting standards and targets and sharing of such information through feedback”. Quality improvement means the “establishment of a program or process to manage change and achieve quality improvement based on performance standards, measurements, and reports”.

Turning Point National Program Office (2002) carried out a cross-nation survey on

performance management based on the four-dimension model in the public health sectors throughout U.S.A. This is the only investigation on overall performance management procedure in the public sectors in recent years, though the survey focuses on how the performance management has been undertaken rather than on its performance-related validity.

Table 3.2 Turning Point’s Performance Management Model

<p>Performance standards</p> <ul style="list-style-type: none"> ● Identify relevant standards ● Select indicators ● Set goals and targets ● Communicate expectation 	<p>Performance measurement</p> <ul style="list-style-type: none"> ● Refine indicators and define measures ● Develop data systems ● Collect data
<p>Reporting of progress</p> <ul style="list-style-type: none"> ● Analyze data ● Feed data back to managers, staff, policy makers, and constituents ● Develop a regular reporting cycle 	<p>Quality improvement process</p> <ul style="list-style-type: none"> ● Use data for decisions to improve policies, programs and outcomes ● Manage changes ● Create a learning organization

Source: Turning Point. From Silos to Systems: Using Performance Management to Improve the Public’s Health (2003).

3.6 Organizational performance

3.6.1 Terminology of the performance

Kates et al. (2000) point out that it is very important to standardize the performance terminology to get a mutual understanding across stakeholders and portray a consistent picture of the performance over time.

Otley (1999) thinks performance refers to the process of doing the job, together with the results achieved. Bernardin (1992) advocates that outcomes should be the “core component” in performance management.

On the contrary, Campbell (1990) maintains that performance is behavior. Others share with the similar idea that the performance management should focus on behaviors rather than outcomes or results that derived from those behaviors (Murphy

and Clevelen, 1991; Latham, 1986)

These two kinds of definitions invoke the debate whether performance encompasses behavior and results, or only one single indicator. Mwita (2000) regards performance as a comprehensive concept that it should circumvent at least three variables: behavior, output, and outcomes. Bernadin et al. (1984) also regard performance as the “record of outcomes produced on a specified job function, activity, or behavior during a specified time period”. In other words, performance is the sum of performance on critical job functions or behaviors.

3.6.2 Private sector performance

The term of ‘performance’ was firstly introduced in the business areas due to the ever-fast changing environment surrounded and the complexity of the organizational structure that could not be tackled through traditional management ways with a purpose to understand and improve performance (Jan Van Ree, 2002).

Jan Van Ree (2002) regards performance as a set of comprehensive indicators, reflecting the behavior of the organization, which could be measured and evaluated. In other words, performance is the quality of the transformation from input to output that would influence the balance sheet of an organization. Scotti (2004) thinks that the “concept of performance deals with how well things are done”.

Generally, performance relates with efficiency and effectiveness. Jan Van Ree (2002) puts forward a more detailed structure of organizational performance model, which consists of five dimensions: effectiveness, efficiency, productivity, flexibility, and creativity.

From a historical perspective, performance structure evolves in accordance with the demands of the environment put on the organization. Weggemen et al. (1992) argue that the constant adding of performance criterion cast light on the challenges and demands that organizations face in the past decades, which also increase the

complexity and difficulty of managing performance as a managerial tool to improve the performance.

3.7 Government performance

3.7.1 Government performance structure

Many scholars have carried out many valuable researches contributing to this agenda.

Barnow (1992) suggests that government performance could be measured with three broad dimensions: gross outcomes, net outputs, and inputs and processes measures. They refer the gross outcomes to the “measures of outcomes of the program at some designated date”, and define net outputs as the “measures of the value added of the program”.

Phillips and Phillips (2004) put forward another kind of performance structure, consisting of three dimensions: effectiveness, efficiency, and impact. Phillips gives definitions on the three dimensions: effectiveness indicates the “comparison of produced output (provided service) to intended output or objectives”. efficiency refers to the extent to that the “government produces a given output with the least possible use of resources”, the impact shows the “macro effects of public transit and reflects the efficiency and effectiveness of transit, as well as external and indirect effects on social well-being, economic development, and environmental quality”.

Hoogenboezem (2004) has the same idea towards the government performance structure. He suggests that performance has two categories, internal indicators and external indicator. The internal indicators measure the organizational output, which is unambiguous and the data is easy to get through calculating a set of ratios between input and output. The external indicators measure “the state of affairs of the larger object of administration”, which is not easy to assess.

Boyne et al. (2003) suggest a criteria model of government performance while

studying the relationship between NPM reform and its consequences. The criteria include responsiveness, efficiency, and equity. Pollitt (2000) shares the same idea, however his criteria are much broader, which comprise “savings, improved processes, improved efficiency, greater effectiveness, and an increase in the overall capacity/flexibility/resilience of the administrative system as a whole”.

Brewer and Huque (2004) outline three dimensions consisting of “service efforts, service accomplishments and relate efforts to accomplishments”, which could be measured in the government sectors. In their words, the service effort includes the amount of both financial and non-financial resources, which are employed so as to provide the service. The accomplishments refer to the ‘outputs’ and ‘outcomes’ of a particular service. The last term of “relate efforts to accomplishments” mainly concentrates on efficiency.

Rogers (1990) introduces a two-orientation model to describe government performance structure: internal/external, employee/institution. He thinks that the combination of the two orientations leads to a way to study government performance. He terms the four kinds of performance structure: employee participation and development (employee/internal), stability and controlling capability (institution/internal), goal setting and output (institution/external), and adaptability and responsiveness (employee/external).

Brooks et al. (2002) put forth an organizational performance framework (ARTD), which includes leadership, customer and market focus, strategy and planning, work environment, work processes, product and service quality, and reputation.

Among all the models of the government performance, the two most famous models are the Fitzgerald’s six-dimension model and the balanced scorecard. Fitzgerald et al. (1991) bring forward a performance model consisting of six dimensions. In this model, the dimensions are divided into two lays: the bottom-layer the determinant factor and the up-layer result factor. The result factor has two dimensions called *primary objects*:

competitiveness and financial success, while the determinant factor, contributing to the success of the results, has four dimensions called *secondary objects*: quality, flexibility, resource utilization and innovation. The relationship between the primary and the secondary objectives is obvious and complementary (Kloot and Martin, 2000).

Kaplan and Norton (1992) develop a managerial tool called balanced scorecard to assess the organizational performance based on Fitzgerald's model, consisting four dimensions: financial, community, internal business processes and innovation and learning. Though this tool was firstly introduced in commercial sectors, it still can be applied to government sectors (Whittaker, 2004). There is an obvious linkage between the two models. Kaplan and Norton integrate Fitzgerald et al.'s three dimensions, quality, flexibility, and resource utilization, into one dimension—internal business processes (Kloot and Martin, 2000).

3.7.2 Balanced scorecard

Kaplan and Norton (1992) set a list of indicators to the four dimensions in the balanced scorecard respectively. The first dimension of financial perspective encompasses revenue growth, cost management and asset utilization. The second dimension of customer perspective circumvents market share, customer retention, customer acquisition, customer satisfaction, and customer profitability. The third dimension of internal process perspective consists of identifying or making the market, designing, building, delivery, and after-sales service. The fourth dimension of innovation and learning perspective makes up of employee capabilities, information capability, motivation and alignment.

Kaplan (2001) makes a conclusion that it is a useful management tool for the non-profit organizations. Chan (2004) maintains that the balanced scorecard is a valuable tool to government sectors though his 2004's survey reveals only a small proportion of the municipality governments in Canada and USA has adopted this

strategic tool. Furthermore, his survey also reveals that most local governments pay much more attention to the financial perspective, least attention to the innovation perspective.

Whittaker (2004) shares with the same idea that the four dimensions used in the commercial field also could be used in the public management area. Furthermore, he makes detailed definitions about the four dimensions in light of the public management means. Financial Perspective relates to “managing the budget to obtain the desired results at minimal cost or with the greatest efficiency”. A combination of saving and reallocation of funds mean returns to government sectors in this regard. Stakeholder Perspective, called as community or customer perspective, comprises the “ability of the organization to provide quality goods and services, effective delivery, and overall stakeholder satisfaction”. Internal Process Perspective focuses on the key operation processes through which government organization could “excel to continue adding value for its stakeholders”. Generally, this perspective includes service development and delivery, and partnership with the stakeholders. Learning & Growth Perspective provides a premise for the other three perspectives to reach their goals in a long run. In the competing world, government heavily relies on multiple competencies of its employees to implement the programs successfully. How to motivate employees, improve their competencies, and build a comfortable organizational climate for sustaining improvements, are vital indicators to this perspective.

3.7.3 Stakeholder theory

In recent years, performance management model has begun to focus on the stakeholders (Mwita, 2000, Kloot and Martin, 2000). For instance, Tat-Kei Ho and Coates (2002) advocate the participation of the citizens during the operation of the government sectors.

Freeman (1984) notes that a stakeholder is a group or individual “who can affect or is

affected by the achievement of the organization's objective". Brooks et al. (2002) regard the stakeholders as those who have "vested interests" in services or programs provided by the sector.

Brooks et al. (2002) argue that the stakeholders should include owners, staff, internal customers, strategic partners, and special-interest groups. Kloot and Martin (2000) outline a list of obvious stakeholders, including the "federal government, ratepayers, the wider local community, consumers and clients of local government services, while employees, councilors and suppliers". Nyhan and Marlowe (1995) refer the stakeholders to citizens, legislators, businessmen, and interest groups, etc. Atkinson et al. (1997) divide the stakeholders into two categories: environmental stakeholders and process stakeholders. The former includes customer, owner and the community, while the latter makes up of employee and supplier.

3.7.4 Monetary versus Non-monetary indicators

Traditionally, government performance measurement aims to construct for an "efficient, effective and accountable government sector" (Guthrie and English, 1997). Consequently, those measures always concentrate on the indicators mainly connecting with "economy (inputs) and efficiency (costs)" (Kloot and Martin, 2000).

Many researchers have criticized the traditional method for the "exclusion of non-financial dimensions of performance" (Ghobadian and Ashworth, 1994, Atkinson et al., 1997, Guthrie and English, 1997). Some scholars even argue that monetary measures could not provide the requisite information for the decision-makers to manage processes (Ballantine et al., 1998; Atkinson et al., 1997; Kaplan and Norton, 1992, 1996; Ghobadian and Ashworth, 1994; Fitzgerald et al., 1991). O'Donovan and Walsh (2000) maintain that the success of the government sectors could not be evaluated by "how close they constrain spending". Smith (1995) argues that the focus of performance indicators on economy and efficiency is dangerous rather than misleading.

Kaplan and Norton's (1992) balanced scorecard methodology argues that organizations should not only focus on the "bottom line"—economic indicators, but balance the "hard" and "soft" measures. They (1996) maintain that performance information is "wider than financial information". Furthermore, they argue the three non-financial dimensions are organization's drivers of "future performance", while the financial dimension only means the past performance.

Phillips and Phillips (2004) argue that the improvements of non-monetary performance indicators "can be converted to monetary value". Moreover, they think that there are intangible benefits apart from those monetary benefits that have been converted into monetary value. Generally, these intangible benefits are often more important to government organization than those monetary benefits.

Kloot and Martin (2000) point out that there are increasingly demand to use these non-financial performance indicators besides the financial measures in the government sectors. Barnow (1992) identifies a range of indicators that have been regularly used in government programs and finds that the "measures do not necessarily correspond to any single economic concept". Lingle and Schiemann (1996) report the government has noticed it is more important of the non-financial measures compared with the financial measures.

3.7.5 Summary

Government performance is the criterion in this study, which means the selection of an appropriate performance structure would determine the success of the study. The balanced scorecard is the most prevalent performance model, both for the private and government sector, throughout the world. Moreover, many field studies have already been undertaken to find the feasibility of introducing it in the government sectors. The results are positive. That is why we choose this model as the theoretical premise for this study. However, the financial perspective will not be included. The reasons have been discussed above, together with the difficulty to obtain detailed and accurate

financial information during the survey.

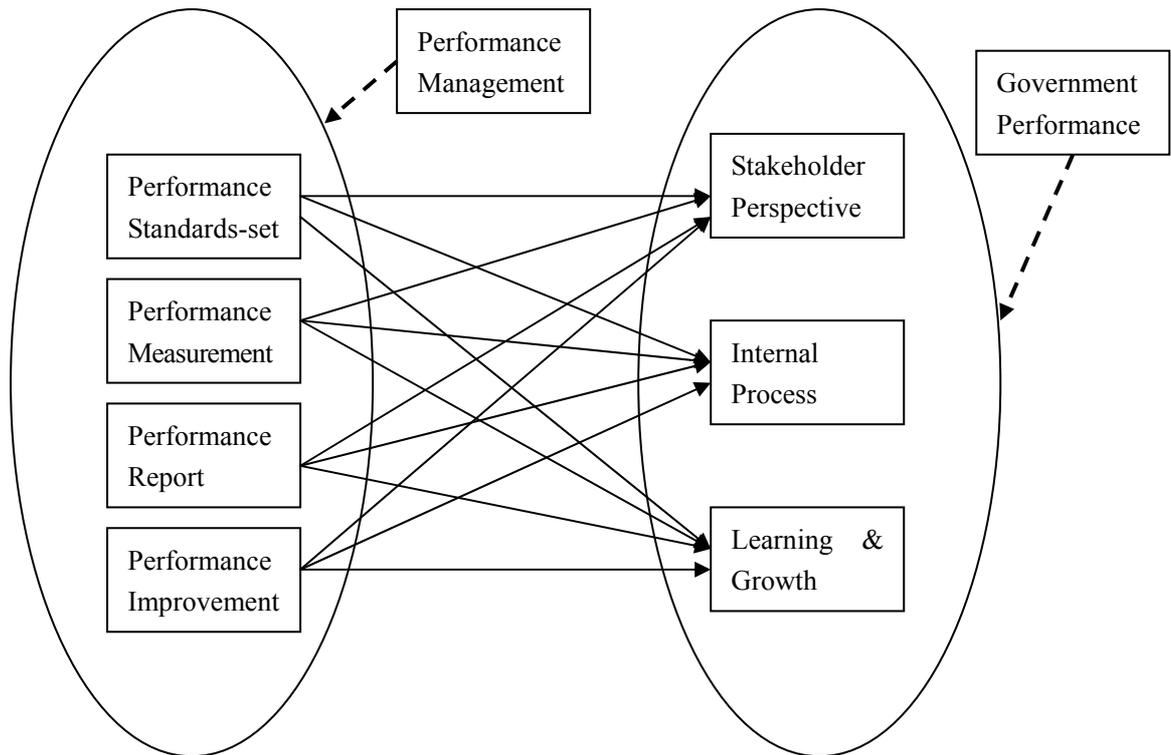
On the other hand, we follow Campbell et al.'s idea that performance is the reflection of behavior. Because this study is focusing on the criterion-related validity in the government sectors as a whole, not on one specific sector, it means that sector-specific performance criteria will not be included in the criteria. Considering the diversity of the government sectors in terms of the multiple goals and functions, it is almost impossible to harmonize these different performance indicators of different government agencies. The only solution is to focus on the universal performance dimensions or measures not specific indicators. By this, the Balanced Scorecard can be used to measure the behavioral results the extent to which the organization has done rather than the specific output of the organization.

3.8 The development of our model

In the theoretical framework, we have discussed the performance management, the independent variable, and the government performance, the dependent variable, with respect to the structure, contents, function and debates in depth. According to the research purpose, the relationship between the two variables and differences, if any, of the ratings on the independent variables between the two nations, Sweden and China, will be the two main focuses. In this way, the structure of the two variables is of vital importance, which is the fundamental factor guaranteeing the reliability and validity of the research.

In light with the research purpose, we put forth a research model, covering two sub-models and their relationships as well. The model is as follows. The detailed interpretation of the model will be discussed later.

Figure 3.1 Research Model



3.8.1 The performance management variable

In the theoretical framework, different kinds of definitions, contents, and models of the performance management have been discussed. The scale has to be based on the theories, especially the discussed models.

We regard the performance management as a set of integrated procedures covering from setting standards, selecting indicators, measuring, reporting, reviewing, and improving, with a clear purpose to improve the government performance so as to improve the accountability as the final target. Furthermore, the procedures should both focus on the strategy and operation.

3.8.1.1 Amendments of the two models

In line with this definition of the performance management from a procedural and strategic perspective, we studied the models have been mentioned closely and found

that the two models, Mwita's and Landrum and Baker's, both need amendments according to the purpose of the study.

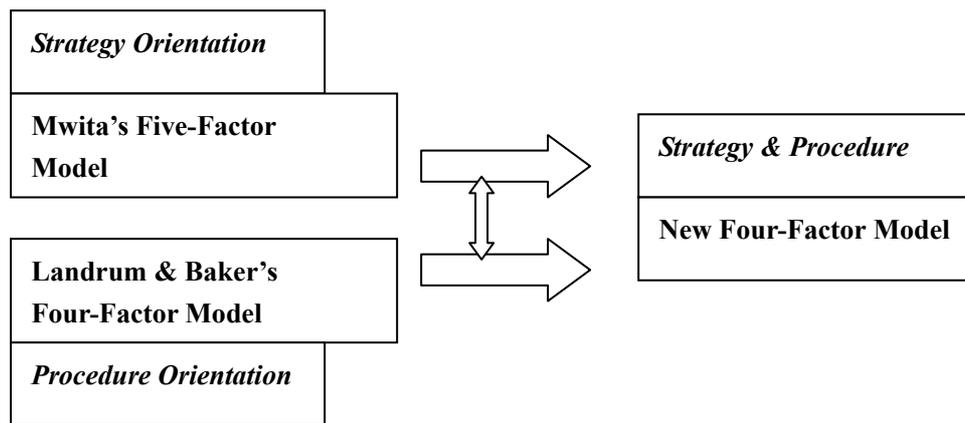
Mwita's (2000) model focuses more on the strategy, while passing over some concrete procedures or steps that the performance management should embody, advocated by many other scholars. Moreover, the first two dimensions, mission statements and strategies and plans, could be to some extent integrated into one dimension from a strategic view. One indicator of the strategies and plans dimension, training and development policies, and one indicator of the performance recognition dimension, training needs assessments, seem to be overlapped or contradictory. Despite these, this model outlines fundamental steps of the performance management from a strategic perspective.

Landrum and Baker's (2004) model has been testified in terms of its feasibility. A little bit different from the Mwita's model, this model focuses more on the solid steps or procedures at operational level. In other words, this model pays little attention to the strategy and stakeholder, which are the two main strands of the performance management, argued by many researchers. Hence, this model should be modified in compliance with the strategic requirements of the performance management.

3.8.1.2 Integrations and adjustments of the models into a new model

We integrate the two models and other researchers' arguments to create a comprehensive model that could embody almost all the steps the performance management should have, focusing on the strategic and operational level simultaneously and harmoniously.

Figure 3.2 Combinations of the Two Models



The new model is designed in compliance with the consecutive procedures of the performance management. From this point, we formulate a four-dimension model based on the Landrum and Baker’s model as the basic framework and integrated with the Mwita’s strategic concepts.

Table 3.3 The New Performance Management Model

<p>I: Performance Standards-Set -Understanding the purpose of existence -Identify relevant standards -Select indicators -Set goals and targets -Communicate expectation in and out the organization -Stakeholder expectation consciousness</p>	<p>II: Performance Measurement -Refine indicators and define measures -Develop data systems -Collect data</p>
<p>III: Performance Report -Analyze data -Feedback/feed forward loops -Develop a regular report and review cycle</p>	<p>IV: Performance Improvement -Use data for decisions to improvement -Manage changes -Rewards and sanctions -Training needs assessment -Create a learning organization -Suggestions box</p>

3.8.1.3 Interpretation of the model

Dimension I: Performance standards-set

The first dimension seeks the information on the establishment of organizational

performance standards, targets, and goals to improve practices. This dimension has several terms that should be explained in detail to ensure that the respondents can understand without any difficulties.

The first term is the **performance standards**, which are the guidelines and always the minimum acceptable outcome that are used to assess organizational performance (www.dnr.state.wi.us/org/water/wm/nps/stormwater/definitions.htm, also Turning Point, 2002). The second term is the **performance indicator**, which is a particular characteristic or dimension used to measure the organizational performance (www.balancedscorecard.org/bkgd/definitions.html). The third term is the **performance target**, which is an intended level of performance expressed as a tangible, measurable objective, against which actual performance can be compared (www.afm.ars.usda.gov/strat98/glossary.htm).

This dimension consists of six indicators. The detailed interpretations are as follows.

The indicator of “understanding the purpose of existence” means that every member should clearly understand the organization’s mission and values, and what it wants to achieve (Mwita, 2002). This indicator just reflects the strategic perspective.

The indicator of “identify relevant standards” means that the organization set standards, the bottom-line, for organizational tasks through different techniques, e.g. benchmarking, so that the task or activity could be evaluated afterwards.

The indicator of “select indicators” means that appropriate indicators should be elaborately chosen so that they could reflect the real performance and be measured easily.

The indicator of “set goals and targets” means that how the organization defines the level it should meet based on the set performance standards. The target is always a little bit higher than the standards.

The indicator of “communicate expectation in and out the organization” means that the organization should proactively communicate the performance standards,

indicators, and targets to the employees and the external stakeholders.

The indicator of “stakeholder expectation consciousness” means that the organization should not regard it as a close system but an open one. Hence, it should consider the stakeholders’ ideas, demands, and suggestions seriously during the operation so as to meet with public expectations.

Dimension II: Performance measurement

This dimension, seeking the information on the development, application, and use of performance measures to assess the actual achievements, has three indicators.

The indicator of “refine indicators and define measures” has two layers meaning. The first is that the organization should scrutinize the performance indicators further so that they could be measured with ease. The second refers to the ways of measuring, which should be defined clearly in advance.

The indicator of “develop data systems” focuses on the establishment of a system concerning on performance data collecting, tracking, and so on.

The indicator of “collect data” refers to the organizational behavior of the performance data collection in line with the set targets.

Dimension III: Performance report

This dimension seeks the information on the documentation and reporting of progress in meeting standards and targets and sharing of such information through feedback. Three indicators comprise this dimension.

The indicator of “analyze data” refers to the behaviors of documenting and analyzing the collected performance data with specific methods.

The indicator of “feedback/feed forward loops” means that the establishment of

feedback/feed forward scheme so that the performance information could be available to the employees and the external stakeholders.

The indicator of “develop a regular report and review cycle” means whether the organization establishes a regular or formal scheme of performance report and review, irrespective of how many times per year.

Dimension IV: Performance improvement

This dimension seeks the information on the establishment of a program or process to manage change and achieve quality improvement based on the performance standards and reports, consisting of six indicators.

The indicator of “use data for decisions to improvement” means that there should be a scheme to use the performance information for performance improvement decision-making and priority setting, resource allocating (Turing Point, 2002).

The indicator of “manage changes” refers to capacity of employees on the organizational development, and the scheme of change facilitation (Turing Point, 2002).

The indicator of “rewards and sanctions” refers to the performance-based compensation system (Mwita, 2004).

The indicator of “training needs assessment” means that the personnel training should be based on training needs analysis in connection with the performance report.

The indicator of “create a learning organization” is rather a broad concept, focusing on achievement culture (Mwita, 2004).

The indicator of “suggestion box” refers to the organizational scheme of proactively collection of improvement suggestions from the employees.

3.8.2 The government performance scale

As to the government performance structure, we choose the balanced scorecard as the theoretical base to design the performance scale. As far as the purpose of our study is concerned, we mainly focus on the general behavioral results of different types of government organizations, which mean organization-specific indicators will not be included. The financial dimension also will not be included; the reasons have been discussed in the theoretical framework chapter together with the difficulty of obtaining accurate and sufficient financial information of so many organizations. According to Kaplan and Norton's (1992) principles, the detailed interpretation of the three dimensions is as follows in line with the research purpose.

Dimension I: Stakeholder Perspective

This dimension focuses on the external stakeholders' satisfaction on the ability of the organization whether it could meet with the demands of the stakeholders. Generally, the stakeholders should rate this dimension. However, considering its difficulty to find appropriate and sufficient representatives of different types of the stakeholders, we use the public employees as substitute of the external stakeholders. This dimension is measured by the satisfaction of the stakeholders in terms of the responsiveness, quality, and timeliness of the goods or services provided by the public agency, numbers of the complaints, cooperation between the agency and the stakeholders, and the impacts of the policy or service on the social and economic development (Phillips, 2004).

Dimension II: Internal Process Perspective

This dimension circumvents the key process excellence of the organization. According to the balanced scorecard, the internal process is divided into three cycles: the innovation cycle, the operation cycle, and the post-sale service cycle. In line with the characteristics of the government agencies, the key processes should include customer identification, service development, service delivery, and post-service. Adherence to these four key processes, this dimension consists of four indicators.

The indicator of “customer identification” refers to the understanding of the customers’ characteristics and demands by the public employees.

The indicator of “program development” is the key point of this dimension, referring to the service development capability, one of the vital processes that could meet with the public demands (Whittaker, 2004).

The indicator of “service delivery” focuses on the efficiency of the service delivery process.

The indicator of “post-service” refers to the follow-up services offered to ensure whether the demands of the customers have been met with.

Dimension III: Innovation and Learning Perspective

This dimension mainly concentrates on the competency of the employees and the way the organization takes to improve it, measured by employees’ satisfaction on four indicators.

The indicator of “empowerment” refers to the satisfaction of employees, especially those low-level, on the empowerment from the leadership of the organization (Olve and Wetter, 1999, Dixon, et al., 1998).

The indicator of “competency of the employee” mainly focuses on the proportion of the competent employees who could match the competitive environment (PEA, 1999).

The indicator of “work climate” refers to the satisfaction of the employees with the atmosphere, underpinning the sustainable development (PEA, 1999).

The indicator of “information capability” focuses on the information availability for decision-making, with an assumption that information is a key point to development (Turning Point, 2002).

3.8.3 Research hypotheses

3.8.3.1 Causal relationships between the independent and dependent variables

The dependent variable has three dimensions: stakeholder perspective, internal process, and learning & growth. The independent variable has four dimensions: performance standards-set, performance measurement, performance report, and performance improvement. The correlation hypotheses are based on the inherent nature of the contents between these dimensions.

The stakeholder perspective reflects the satisfaction of the stakeholders. According to the contents, two dimensions of the independent—performance standards-set and performance report—have something related to the stakeholder. The performance standards-set dimension requires the organization to take the stakeholder into consideration during the standard setting process, while the performance report dimension advocates the organization to release the performance information to the stakeholders, internal and external. These two kinds of tactics will have positive impact on the stakeholder satisfaction. Thus, we come to the first hypothesis.

H₁: There are significant causal relationships between two dimensions of the independent variable—performance standards-set and performance report—are and the stakeholder perspective.

The internal process mainly focuses on the efficiency of the operational processes. There is one dimension from the independent variable—performance improvement—that is concentrating on the efficiency and effectiveness enhancement. In other words, performance improvement may lead to efficiency improvement, the key component of the internal process. We then come to the second hypothesis.

H₂: There are significant causal relationships between one dimension of the independent variable—performance improvement—and the internal process.

The learning and growth circumvents the competency of the employee and the sustainable culture within the organization. The performance improvement dimension relates to development and train. From the contents of the two dimensions, we could see some conjunctions between them by nature. We present the third hypothesis.

H₃: There is significant causal relationship between one dimension of the independent variable—performance improvement—and the learning and growth.

3.8.3.2 Differences between agency levels

As far as the government hierarchy is concerned, Chinese local governments at the three different levels are in charge of different kinds functions towards the public and society. Generally, the municipality government focuses on macro functions, e.g. regulation making and disputes resolution, the district/county government shall carry out the directives from the municipality government in its territory together with limited regulation making, while the sub-district/town government mainly implement administrative directives from its superior governments for the interest of citizens in its small territory. In other words, the sub-district/town government has a very close linkage with the public while the municipality government only establishes comparatively loose ties with the citizen.

From this point of view, performance report may receive different attention between the governments at three levels. The district/county and sub-district/town government may highlight this factor much more than the municipality government due to their close connection with the citizen.

H₄: One dimension of the independent variable—performance report— significantly differentiates between the government agencies at three levels.

With regard to the government performance, it is the same case for the stakeholder perspective. Closer connection may lead to higher rating on this factor.

H₅: One dimension of the dependent variable—stakeholder perspective— significantly differentiates between the government agencies at three levels.

3.8.3.3 Differences between two nations

The introduction of the New Public Management concepts and performance reforms in Sweden is much earlier than China. Therefore, the Swedish local government has gained experience in this regard, which means that Sweden should be better than China in terms of the performance management procedures.

H₆: The ratings on the four factors of the performance management are significantly different between Sweden and China.

With regard to the government performance, there also should be some differences between the two nations. It is, however, quite hard to figure out where the difference lies. In the past decade, Chinese government focuses greatly on the demands of the people and the administrative efficiency of the government agencies, which means that there are no significant differences on the two factors—stakeholder perspective and internal process—between China and Sweden. Compared with the Swedish government, the Chinese government pays much attention to the short/middle-term interests. Long-term strategy is not always the top priority to the leadership. Therefore, difference may lie in the third factor of learning and growth.

H₇: The rating on the learning and growth is significantly different between Sweden and China.

Chapter 4 Empirical Method

The empirical method is presented. The research strategy, focusing on the way we implement the research, is described. The samples in terms of the way of selection process and general information are discussed. Limitations of the survey are also described. Further, the questionnaire is discussed. The chapter is ended with the validity, reliability, and generalization analysis.

4.1 Research Strategy

We created our model based on an extensive literature review and put forward hypotheses focusing on the relationship between the independent and dependent variables, and on differences, if any, between the two nations on the ratings on the four dimensions of performance management.

To prove these hypotheses, we conducted an empirical study. Considering the characteristics of the study, we chose questionnaire as our research tool. Firstly, a questionnaire can cover a larger range of population than other means like an interview. Secondly, a questionnaire is a feasible way to carry out an international comparison study. Thirdly, the Likert scale is a suitable tool used to collect useful data for statistical analysis. In this study, we designed a self-assessment questionnaire for its convenience of implementation and control.

Before the test, we made a pilot test to refine the items, ensuring that the respondents have no difficulty in understanding these items. This is a useful way to ensure the validity and reliability of the questionnaire (Saunders, 2003, 308). The English version was mailed to some teachers from different departments of the Kristianstad University. We regard the teachers as good subjects for the pilot test for the congruence between them and the target subject, the public servants. According to the research, we will survey the employees of the government agency, one kind public

sector. On the other hand, a school is a typical public sector, sharing considerable similarities with the government agency. Thus, the pilot test on the teachers can be proved reasonable. Feedback information was deliberately analyzed and amendments were made accordingly.

Meanwhile, the Chinese version was sent to our classmates, who are managers the government agencies, and our friends in Chinese government agencies to make a pilot test. Amendments on the Chinese version were made based on that feedback.

After the pilot test, we implemented the survey in dual ways, online and paper-pen methods. The survey in Sweden was carried out solely online, while the majority of the Chinese data was collected by using the paper-pen method. A survey online is the most efficient way to collect as much more information as possible in a short period time. There are some other explicit advantages related to the online survey. First, it is convenient for the respondents to answer the questions online. Second, it makes large-scale survey possible, especially in two distant countries. Third, survey online could guarantee the respondents that their personal answers would be kept confidential so as to motivate them to make objective evaluations on the performance of his/her own organization.

The questionnaire was put online with detailed interpretation of the purpose and how to finish the survey. The website of the English version is at www.dx315.gov.cn/vote/default_english.htm, while the Chinese version locates at www.dx315.gov.cn/vote/default_chiense.htm. A specific computer program was also designed, making it possible for us to get the result information immediately after the respondents have done the survey.

Noticing the low response rate with respect to the online survey (Coomber, 1997), we called our colleagues and friends in government agencies in Ningbo municipality, asking them to motivate their colleagues to participate in the survey online. In Sweden, our program assistant Lisa helped us write a letter in Swedish, making a brief

introduction of the survey and us. Two hundred emails have been sent to directors and employees in different government agencies in three municipalities, in Kristianstad, Osby, and Hässleholm. As to the paper-pen method, we asked our friends to make paper copies and send to their colleagues to answer, ensuring more people can participate in the survey.

4.2 Sample

This research only focuses on local government agencies, not wider concept—local public sectors. So the definition of the different organizations and their clear distinction is of vital importance in the survey, otherwise unclear data from different kinds of organizations will distort the results.

Government agency means administrative unit of government. While *administrative agency* is a governmental body responsible for administering and implementing a particular legislation, such as laws governing workers' compensation. These agencies may have rulemaking power and judge-like authority to decide disputes. On the other hand, *public sector* is a much wider concept, which means part of the economy concerned with providing basic government services. In most countries the public sector includes such services as the police, military, public roads, public transport, primary education and healthcare for the poor.

In line with these definitions, we chose our subjects strictly. After analyzing the governmental structure in Sweden, only administrative agencies under the municipality executive committee are included in the survey. Other public organizations affiliated to administrative agencies were included in the survey if they were simply providing services. This procedure was also the case while selecting the subject agencies in China.

All the Chinese respondents are from sixty-five government agencies at three levels, municipality, district/county, and sub-district/town in the Ningbo municipality. The Swedish respondents are from five agencies in three municipalities.

4.3 Limitations

As far as the way of the implementation of the survey is concerned, there are some limitations, which should be taken into account while analyzing and generalizing the validity of the results.

Firstly, there is no further segmentation of the agencies participating in the survey. Although the performance indicators specifically focus on managerial behavior not specific output, performance dimensions in different types of the agencies may have different embodiments, which may hamper the validity generalization on government agencies as a whole. Secondly, the performance scale is rated by the public employees, which may not show the real picture of performance, especially the stakeholder dimension. Thirdly, online survey highly depends on the credibility of the respondents. Thus, quality control of the survey becomes very difficult or impossible. Fourthly, the translation of the questionnaire may bring or discard information of the original version, which maybe distort the comparison.

4.4 The questionnaire

The questionnaire has two scales, a performance management scale, and a performance scale, with a total of 40 questions. The questions are designed according to Likert's 7-point scale, ranging from 7 of "almost always true" to 1 of "almost never true".

The respondents are asked to evaluate the current practices of performance management and organizational performance objectively, based on the real situation in their own organizations. This kind of self-assessment questionnaire has been widely used in the management field. Moreover, Watson et al. (1991) note that internal performance measurement could reflect actual performance better than external measurements of customer satisfaction.

In the first part of the questionnaire (appendix 1, 2), there are several lines of instruction on the purpose of the survey and how to answer the questions. This is a good way to make the respondents understand the survey better and answer the questions more easily.

Before the questions, the respondents are asked to provide information about themselves and their organization. Information about, gender, age, career years, organizational size is focused in this part.

The subsequent questions, the performance management scale has four sub-scales: performance standards-set, performance measurement, performance report, and performance improvement. The total questions of this scale are 25. Questions from 1 to 8 are designed to focus on the performance standards-set dimension. Questions from 9 to 12 concentrate on the performance measurement dimension. Questions from 13 to 17 circumvent the performance report dimension. Questions from 17 to 25 seek information on the performance improvement dimension.

The government performance scale follows behind the performance management scale with three sub-scales: stakeholder perspective, internal process, and learning and growth. The total questions of this scale are 15. Questions from 26 to 31 cover the stakeholder perspective. Questions from 32 to 36 concentrate on the internal process. Questions from 37 to 40 circumvent the learning and growth.

Each sub-scale has a heading followed by a detailed explanation about what kind of information it wants to collect, together with interpretations of some tricky terms to guide the respondents.

4.5 Response rate

It is a little bit difficult to calculate the actual response rate due to the use of both an online survey and a paper survey. As far as the online survey is concerned, it is very hard to know how many people have been asked to participate in the survey. In

Sweden, we emailed not only to employees but also to director, which makes it very hard to know how many employees have been informed. If we only take the sent emails (200) into account, the superficial response rate is 5% (10/200). In China, we asked our Chinese friends to motivate their colleagues or subordinates to join in the survey, however, the number of the people they have asked is not clear. With respect to the paper survey, it is quite easy to figure out the response rate. Nevertheless, this response rate cannot show the whole picture. From this point of view, we do not persist to calculate the response rate.

4.6 Reliability

Reliability refers to “consistency, stability, or the repeatability of research” (Christiansen 2004, 182). Robson (2002) argues that four factors may affect the reliability. The first is participant error. Selecting a neutral or “correct” time to send questionnaire can reduce the possibility of participant error (Saunders, 2003, 101). For instance, research conducted on different days within a week may have different results. We sent emails and made calls on Tuesday, ensuring that the respondents may concentrate themselves on the survey. The second threat is participant bias. The self-assessment may reflect the extent to which the management wants the employees to say (Saunders, 2003, 101). Consequently, the assessment results are distorted. To reduce this possibility, we deliberately chose to use an online and anonymous survey to motivate the respondent’s objective evaluation on the reality. As to observer error and bias, we can avoid it because we only use a questionnaire, consisting of questions that are only explained to one meaning.

4.7 Validity

Validity examines the extent to which the finding really accords with the reality, and finds whether there is causal relationship between two variables (Saunders, 2003, 101). Generally, the higher the quality of the research is, the higher the validity should be. But, it is difficult for researchers to ensure the validity of the research, because the

questions used in the survey are very likely to express misleading meanings. Furthermore, the respondents are also likely to answer untruthfully. In this regard, questionnaire design process is of vital importance. The questions of two sub-scales in our questionnaire, one on performance management, and the other on government performance, are deliberately designed strictly based on theoretical models. Moreover, these questions are operationalised into quantifiable statements. Meanwhile, before making the survey, we have sought suggestions from specialists in the public management area and the English language; the questionnaire amendment was revised based on their comments. So, we think both the construct and content validity of our research are high.

4.8 Validity generalization

The problem with generalization means whether the results of this study could be generalized to a larger scale of population (Saunders, 2003, 102). The biggest challenge to this effort is whether the selected sample is of sufficient numerical sizes (Saunders, 2003, 87).

As to this study, the effort of the validity generalization is confined to the Ningbo territory. Considering the social and economic development level, and government structure, China is not a uniform country in this regard. From this point, the sample of Ningbo, a rather developed municipality compared with most parts in China, can only represent some developed provinces and cities in China, mainly in the southeastern provinces, which limits the validity generalization.

Furthermore, the “perception of the respondents and a potential of self-selection bias” related to the online survey may affect the validity generalization (Chan, 2004).

Chapter 5 Analysis of the Survey

The survey is analyzed. We present the results of the questionnaire and testify the hypotheses.

5.1 Introduction of the statistical methods

In this dissertation, statistics is the fundamental way to confirm the hypotheses we have put forth, and to find out specific phenomena with reference to performance management. Six types of statistical methods are introduced.

(1) Factor analysis. The construct validity of the research model must be verified at the first step, or the following research based on the model will be null. Factor analysis will be used to verify whether the structure of the two sub-models, the performance management and the government performance, is consistent with the theory.

(2) Correlation. Correlation analysis is aiming to testify whether there are significant correlations between the independent and dependent variables so as to confirm the hypotheses. Pearson correlation is the major method.

(3) Regression. Regression analysis is employed to establish a mathematical model of the relationship between the independent and dependent variables. Linear regression with stepwise method will be chosen.

(4) Mean comparison. To know the differences between different government levels in terms of seven dimensions of the research model, mean comparison method is the only choice. ANOVA or T test are the options. If the group number is only two, these two methods are the same. While the number of groups is above two, only ANOVA is feasible.

(5) *Descriptive statistics.* This kind of statistics is aiming to present a general picture of the data. Several parameters will be adopted, e.g. mean, SD, etc.

(6) *Nonparametric statistics.* The small Swedish sample must be treated with nonparametric statistics. With regard to the mean comparison, the Mann-Whitney method is employed in this study.

5.2 Components of the respondents

There are two samples, one from Sweden, and the other from China.

Considering the fact that only ten Swedish respondents, from five agencies in three municipalities, participated in the survey, a few lines are needed make the general picture of this sample. From the age perspective, 50% of the respondents are older than fifty; the others are between 40 and 50. Seven of them are male. Sixty percent of the respondents have worked more than twenty years, while the number of years in the current agency is quite evenly distributed. Fifty per cent of the respondents have a master degree. Top and middle managers account for sixty percent. Their agency size is distributed quite evenly.

As far as the Chinese respondents are concerned, one hundred and ninety-six Chinese respondents from different departments took part in our survey. To make a general understanding of these participants, tables from different angles are used.

Personal information data

Seven dimensions are used dealing with the personal information: age, gender, career years, years in the current agency, job rank, and educational level, and agency level.

Table 5.1 Distribution of Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	>50	9	4.6	4.6	4.6

40-50	38	19.4	19.4	24.0
30-40	93	47.4	47.4	71.4
20-30	56	28.6	28.6	100.0
Total	196	100.0	100.0	

Table 5.2 Distribution of Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	117	59.7	59.7	59.7
	Female	79	40.3	40.3	100.0
	Total	196	100.0	100.0	

Table 5.3 Distribution of Career Year

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-5	43	21.9	21.9	21.9
	5-10	46	23.5	23.5	45.4
	10-15	51	26.0	26.0	71.4
	15-20	22	11.2	11.2	82.7
	>20	34	17.3	17.3	100.0
	Total	196	100.0	100.0	

Table 5.4 Distribution of Year in Current Agency

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-5	75	38.3	38.3	38.3
	5-10	48	24.5	24.5	62.8
	10-15	40	20.4	20.4	83.2
	15-20	16	8.2	8.2	91.3
	>20	17	8.7	8.7	100.0
	Total	196	100.0	100.0	

Table 5.5 Distribution of Educational Level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Master or higher	34	17.3	17.3	17.3
	Bachelor	130	66.3	66.3	83.7
	University studies	32	16.3	16.3	100.0
	Total	196	100.0	100.0	

Table 5.6 Distribution of Job Rank

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Top	13	6.6	6.6	6.6
	Middle	70	35.7	35.7	42.3
	First-line	113	57.7	57.7	100.0
	Total	196	100.0	100.0	

Table 5.7 Distribution of Agency Level

Municipality Level (N=96)	Ningbo administrative bureau of industry and commerce
	Ningbo finance bureau
	Ningbo environmental protection bureau
	Ningbo tourism bureau
	Ningbo labour and social security bureau
	Ningbo personnel bureau
	Ningbo communication bureau
	Ningbo local taxation bureau
	Ningbo state taxation bureau
District/County Level (N=60)	Jiangdong district personnel bureau
	Jiangdong district trade bureau
	Jiangdong district labor and social security bureau
	Jiangdong district health bureau
	Haishu district construction bureau
	Haishu district finance bureau
	Beilun district economy development bureau
	Beilun district personnel bureau
	Daxie Island development district personnel bureau
	High technology park district economy development bureau
	High technology park district administrative bureau of industry and commerce
	Yingzhou district administrative bureau of industry and commerce
	Ninghai county foreign trade bureau
Sub-district/town Level (N=40)	Beizhang sub-district, Jiangdong district
	Dongjiao sub-district, Jiangdong district
	Dongsheng sub-district, Jiangdong district
	Yongjiang sub-district, Jiangbei district
	Beisha sub-district, Jiangbei district

	Hongtang town, Jiangbei district
	Yuehu sub-district, Haishu district
	Lingta sub-district, Haishu district
	Ximen sub-district, Haishu district
	Nanmen sub-district, Haishu district
	Longguan town, Yinzhou district
	Songao town, Fenghua county
	Dongchen town, Xiangshan county

From these seven tables, it is very easy to portray a general picture of the respondents. From the age perspective, the majority—over 80%—of the respondents are younger than forty. Male respondents are more than the female respondents. The career year and years in the current agency are evenly distributed among the respondents. With respect to the educational level, all the respondents have a university experience, which means all the respondents have the lingual capability to understand the questions in the questionnaire. As far as the job rank distribution is concerned, three levels, from top to first-line, are all well represented.

With regards to the distribution of the three agency levels, the respondents are spread over 35 agencies, 9 agencies from the municipality level, 13 agencies from the district/county level, and 13 agencies from the sub-district/town level. All in all, the sample is of high quality regarding the representative since all levels are well represented.

5.3 Factor analysis results

The factor analysis is based on the Chinese data.

The two scales, the performance management scale and the government performance scale, are developed based on theories, discussed in chapter 3. The performance management model is established based on the combination of two models, which means verification is indispensable. The government performance model is designed according to the balanced scorecard with limited clues. In this regard, it is also of vital importance to testify the structure whether there are actually three perspectives with

regards to the performance in a government agency. Under such circumstances, factor analysis is introduced to legitimize the two models and to pave the way for the following calculations such as correlations, regressions, and means comparisons, which are all based on the reliable structure of the two models. In other words, it is a way to verify the construct or theory validity of the questionnaire.

5.3.1 Factor analysis results of the performance management scale

Firstly, we use Kaiser’s standard (Eigenvalue>1) together with the scree plot to determine the number of the factors. Secondly, we choose the principal component as the extraction method. Thirdly, the varimax rotation method is adopted.

5.3.1.1 Number of the factors

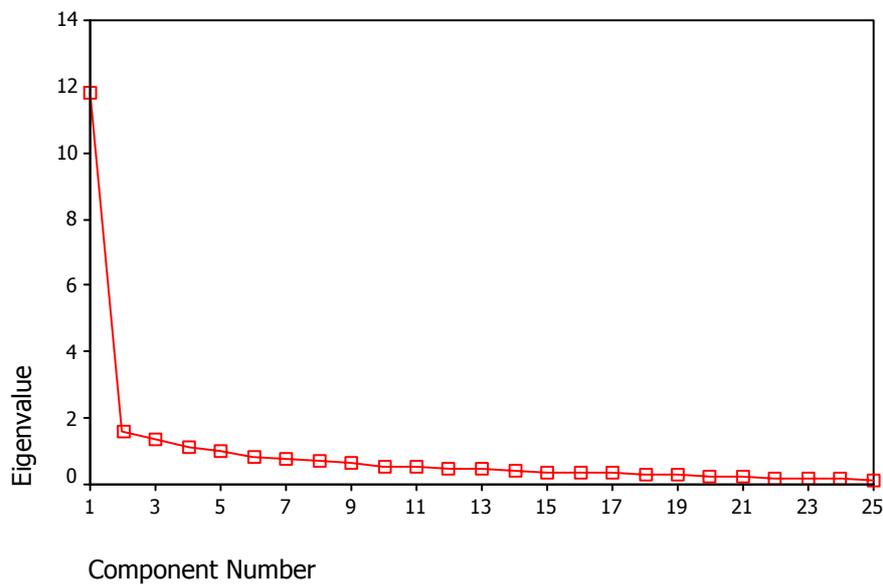
The results reflect a high validity and Chinese specific characteristics with respect to government operation as well. The use of Kaiser’s standard and scree plot both show that there are really four factors related to the performance management model. The four factors can explain 63.5% of the total variance, which is regarded as a high level.

Table 5.8 Kaiser’s standard

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	11.800	47.201	47.201
2	1.579	6.314	53.515
3	1.370	5.482	58.996
4	1.128	4.510	63.507
5	.979	3.917	67.423

Chart 5.1: Scree Plot

Performance management



5.3.1.2 Matrix of the four factors

After rotation, all the items on the performance management scale are clustered into four factors. The results show that most of the items converge consistently with the theoretical hypotheses, with an exception that the performance measurement factor is broadened to a larger perspective combining with several aspects from the original performance report factor.

5.3.1.2.1 Performance standards-set factor

Table 5.9 Rotated Component Matrix of Performance Standards-set Factor

No.	Item	Loading	α
3	The agency identifies performance standards through comparing with historical performance or similar agencies' performance.	.710	.8587
4	The agency selects appropriate performance indicators (characteristics or dimensions, e.g. customer satisfaction) to assess agency's performance.	.646	

2	The agency establishes performance standards (minimum acceptable outcome) for every organizational activity.	.637
6	The performance standards, indicators, and targets are communicated within the agency to ensure that all the employees understand them.	.603
1	All the employees are familiar with the agency's mission towards the public.	.586
5	The agency sets performance targets (intended objective) for activities to be achieved in a certain time period.	.568
8	The agency takes stakeholders (e.g. employees, superior authority, taxpayers, consumers, councilors and suppliers) into account when setting the performance standards.	.533

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 11 iterations.

Table 5.9 shows the convergence of seven items of the original performance standards-set factor. It is in conformity with the theory. However, item 7 does not appear in the matrix. The table also reveals that the Cronbach's α coefficient is very high (.8587), which means that the reliability of this subscale is very high in terms of internal consistency.

5.3.1.2.2 Performance appraisal factor

Table 5.10 Rotated Component Matrix of Performance Appraisal Factor

No.	Item	Loading	α
12	The agency establishes a system for collecting and tracking the agency performance data.	.751	.9083
10	The agency clearly defines ways (e.g. benchmarking, etc.) how to measure organizational activity performance.	.740	
11	The agency establishes a system for collecting and tracking the agency performance data.	.704	
13	The agency documents progress related to performance standards and targets.	.636	
9	Agency takes means to make the performance indicators (characteristics or dimensions, e.g. customer satisfaction) simple, transparent and precisely reflect the actual performance.	.616	

17	The agency performance data are analyzed and reviewed according to set performance standards, indicators, and targets.	.547
14	The agency has a system that regularly (e.g. monthly, or quarterly, or annually) reports performance data from activities and divisions.	.507

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 11 iterations.

This factor matrix reflects a combination of two factors in the research model. Item 9, 10, 11, and 12 all belong to the original factor of performance measurement. However, item 13, 14, and 17 are from the original factor of the performance report. We study these three “new” items closely and find that they are relating to performance record and review. The combination of the two parts reflects the actual managerial practices in Chinese local government. In recent years, performance appraisal has been one of the most important managerial tools to improve accountability and performance in local government on different levels. This tool is rather a broad concept combining measurement, analysis and evaluation. Therefore, this new factor is a reflection of the Chinese characteristic in terms of government performance management tool. Considering the combination of the two parts, we rename the new factor “performance appraisal”. The internal consistency coefficient is also very high (.9083), guaranteeing the reliability of this new subscale.

5.3.1.2.3 Performance report factor

Table 5.11 Rotated Component Matrix of Performance Report Factor

No.	Item	Loading	α
16	The agency regularly reports the agency’s performance information to external stakeholders (e.g.: superior authority, taxpayers, consumers, councilors and suppliers).	.793	.8354
7	The agency regularly reports the performance standards, indicators, and targets to the external stakeholders (e.g.: superior authority, taxpayers, consumers, councilors and suppliers).	.725	

15	The agency performance data are available to the employees through feedback.	.514
----	--	------

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 11 iterations.

This factor is rather small regarding the number of the items. Compared with the original performance report factor, the new factor consists of fewer and mixed items. Item 15 and 16 stems from the original factor of performance report, while item 7 comes from the original factor of performance standards-set. After scrutinizing these three items, they all involve communicating performance information with internal and external stakeholders. In other words, Chinese public servants confine the performance report to a different concept compared with the research model. On the other hand, the contents of this factor is not only relating to the performance data release but also the performance-related information (e.g. performance standard, indicator, etc.) release. Considering its communication feature, the name of performance report is still preserved. As far as the reliability of this subscale is concerned, the internal consistency coefficient is quite high (.8354).

5.3.1.2.4 Performance improvement factor

Table 5.12 Rotated Component Matrix of Performance Improvement Factor

No.	Item	Loading	α
20	The staff has the authority to make certain changes (e.g. change work procedure) to improve performance.	.706	.8801
24	The analysis of employee training needs is based on the performance report	.696	
22	Employees are encouraged to use new ways to improve performance.	.597	
25	The agency has a specific way to collect suggestions on performance improvement from the employees.	.570	
23	Rewards and sanctions are related to performance.	.547	
18	Performance reports are effectively used for decision-making.	.528	
21	The agency has specific means to cope with changes with regard to policies, programs, or organizational structure.	.526	

19	Agency performance information is used to set priorities.	.384
----	---	------

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 11 iterations.

This table shows that the construct validity is very high, because the items from the original factor of performance improvement in the theoretical model converge into one factor. Only item 19 have a comparatively low factor loading (.384). The Cronbach's internal consistency coefficient is very high (.8801)

5.3.1.3 Summary

The factor analysis shows that the construct validity of the performance management scale is quite high. Four subscales all enjoy very high reliability in terms of internal consistency. Compared with the research model, two factors—performance standards-set and performance improvement—are strictly in conformity with the theory. Two other factors reassemble partly; the original factor of performance report is divided into two parts. One part is combined with items from another factor of performance measurement and creates a new factor, which is renamed “performance appraisal” reflecting the current managerial practices in Chinese local government agencies. The other factor, “performance report”, is a comparatively small conception confined to communication with stakeholders.

The results of this part underpin the following statistics. Factor scores are calculated based on the reassembly of the four factors.

5.3.2 Factor analysis of government performance scale

We use both Kaiser's standard (Eigenvalue>1) and the scree plot to determine the factor number. Then, we choose the maximum likelihood method varimax rotation method to yield the factor matrix.

5.3.2.1 Numbers of the factor

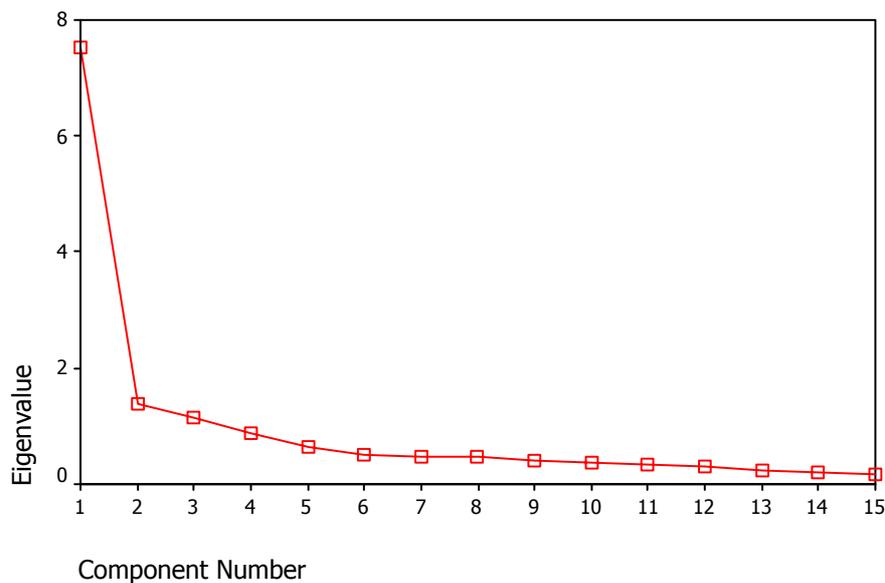
The results reflect high validity on this scale. Kaiser's standard and scree plot all show that there are three factors related to the government performance. The three factors can explain 66.9% of the total variance, which is regarded as a high level.

Table 5.13 Kaiser's standard

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	7.531	50.205	50.205
2	1.370	9.131	59.335
3	1.137	7.579	66.914
4	.884	5.892	72.806

Chart 5.2 Scree Plot

Government Performance



5.3.2.2 Matrix of the three factors

The factors analysis shows that all items converge consistently with the research model; with an exception of item 31, the effect of the agency's policy and services from the original factor of stakeholder perspective reassembles into another factor.

5.3.2.2.1 Stakeholder perspective

Table 5.14 Rotated Component Matrix of Stakeholder Perspective

No.	Item	Loading	α
26	Customers are satisfied with the responsiveness of the agency in terms of providing goods or services.	.815	
27	Customers are satisfied with the quality of goods or services.	.787	
28	Customers are satisfied with the timeliness of goods and services.	.734	.9014
30	Stakeholders are satisfied with the agency's cooperation with them.	.599	
29	Complaint cases of the stakeholders towards the agency are very rare.	.489	

Extraction Method: Maximum Likelihood. Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 7 iterations.

Five items of the original factor of stakeholder perspective converge together in accordance with theory. Cronbach's internal consistency coefficient is .9014, which is considered a very high level.

5.3.2.2.2 Internal process

Table 5.15 Rotated Component Matrix of Internal Process

No.	Item	Loading	α
33	The agency can identify the demands of customers in advance.	.657	
34	The agency's service program development is effective.	.624	
31	The agency's policy or service has a very positive impact on local social and economic development.	.611	.8333
35	The service can be efficiently delivered to the customers once the demands are recognized.	.608	
32	The agency clearly knows whom it serves.	.603	
36	The agency's follow-up service effectively guarantees the customers' demands to be fully met.	.406	

Extraction Method: Maximum Likelihood. Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 7 iterations.

All the items from the factor of internal process of theoretical model converge together, together with another outside item (31) stemming from the original factor of

stakeholder perspective. This item, mainly focusing on the outcome of the performance, is heterogeneous to other items of internal process. Therefore, this item is discarded in the following statistics. This exclusion reflects a small flaw of the research model with respect to the stakeholder perspective. On the other hand, the convergence of all the items from the original factor of internal process reflects high construct validity of this subscale. As to the reliability in terms of internal consistency, it is quite high (.8333).

5.3.2.2.3 Learning and Growth

Table 5.16 Rotated Component Matrix of Learning and Growth

No.	Item	Loading	α
38	All the employees have the capability to match the demanding environment.	.711	
40	Reliable information is available for strategic decision-making.	.695	.8416
39	Employees are satisfied with the work climate.	.679	
37	Low-level employees (e.g. assistant, secretary) are satisfied with the empowerment.	.647	

Extraction Method: Maximum Likelihood. Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 7 iterations.

The four items of the factor of learning and Growth converge consistently with the theoretical model, showing high construct validity. Furthermore, the internal consistency coefficient is .8416, manifesting high reliability of this subscale.

5.3.2.3 Summary

The results of the factor analysis indicate a high construct validity of the scale, with an exception of item 31 originally stemming from the stakeholder perspective. The three subscales all show a very high reliability regarding the internal consistency.

5.4 Descriptive analysis

5.4.1 General descriptive analysis

This part is also based on the Chinese data.

The descriptive analysis gives a general picture of the factor scores of the seven factors respectively. With regards to the factor scores, we use a mean of all the items of one individual factor to calculate the scores, disregarding weighted loadings. The table 5.17 reveals the most important parameters of descriptive analysis: mean, SD, minimum, and maximum, and together they portray a clear picture of the seven factors.

When studying the four factors of the performance management scale, we find that the mean of the first factor—performance standards-set—is quite high, while the fourth factor—performance improvement—has a comparatively low score. The other two factors share similar and moderate scores. The rating scores are also quite accordant with the actual situation in Chinese local government. In recent years, a management reform labeled “Object Management Appraisal” prevails in local governments at all levels, with a main feature of clear performance standards and indicators setting at agency level. However, sometimes this reform degenerates into formalism, which means the first step—standards-set—is always performed nicely while the following three steps to some extent are neglected inadvertently, especially the ultimate practice—performance improvement—aiming to improve performance and accountability based on actual performance appraisal. This is a feasible reason why the scores differentiate across the four factors.

When turning to the performance scale, we find that the rating of the learning and growth factor is lower than that of the other two factors, revealing the current situation in the government agencies that the issue of long-term and sustainable development has not yet been received much attention.

Table 5.17 Descriptive Statistics

	Mean	Minimum	Maximum	Std. Deviation
Performance standards-set	4.9694	1.71	6.71	1.02675
Performance appraisal	4.6589	1.00	6.86	1.15318
Performance report	4.6378	1.00	7.00	1.28934
Performance improvement	4.4005	1.25	6.25	1.08633
Stakeholder perspective	5.1224	1.40	7.00	1.00118
Internal process	5.1622	2.20	7.00	.97560
Learning and Growth	4.7781	1.50	6.75	1.16238

5.4.2 Differences of the ratings between job ranks

Coming from different job levels, the respondents may have different motivation with regard to the rating of performance-related issues of their own agency. For instance, politics may have a greater impact on the top managers to increase the ratings. Therefore, the next step of the descriptive analysis is trying to find out whether there are any differences of the ratings of seven factors between the different job ranks. The ANOVA results show that there are no significant differences between rating of the managers and first-line employees. In this regard, the quality of the data is guaranteed after confirming that the job ranks have little impact on the ratings.

Table 5.18 ANOVA

		Sum of Squares	Df	Mean Square	F	Sig.
Performance standards-set	Between Groups	2.312	2	1.156	1.097	.336
	Within Groups	203.260	193	1.053		
	Total	205.571	195			
Performance appraisal	Between Groups	1.694	2	.847	.635	.531
	Within Groups	257.623	193	1.335		
	Total	259.317	195			
Performance report	Between Groups	.174	2	.087	.052	.950
	Within Groups	323.995	193	1.679		
	Total	324.170	195			
Performance improvement	Between Groups	2.420	2	1.210	1.026	.361
	Within Groups	227.703	193	1.180		
	Total	230.122	195			
Stakeholder	Between Groups	.152	2	.076	.075	.928

perspective	Within Groups	195.309	193	1.012		
	Total	195.461	195			
Internal process	Between Groups	4.612	2	2.306	2.459	.088
	Within Groups	180.989	193	.938		
Learning and Growth	Total	185.601	195			
	Between Groups	2.688	2	1.344	.995	.372
	Within Groups	260.782	193	1.351		
	Total	263.471	195			

5.5 Correlation and Regression analysis

The first three theoretical hypotheses put forth in the end of the chapter 3 all circumvent the relationship between the independent and dependent variables. The optimal statistical methods are correlation and regression. The former can reveal whether there are significant correlations between variables, while the latter can testify the existence of a causal relationship. The correlation and regression analyses in our study are also based on the Chinese data.

5.5.1 Correlations between independent and dependent variables

Table 5.19 is the Pearson correlation matrix between the two types of variables, indicating very significant correlations between the four performance management factors and the three government performance factors. It seems that the first three hypotheses are confirmed, however, correlations cannot testify the causal relationship between them and comparative importance of the independent variable to the dependent variable. Thus, the regression method must be taken into consideration.

Table 5.19 Correlations Matrix between Independent and Dependent Variables

	Stakeholder perspective	Internal process	Learning and Growth
Performance standards-set	.511**	.564**	.569**
Performance appraisal	.487**	.567**	.522**
Performance report	.384**	.419**	.487**
Performance improvement	.548**	.624**	.669**

5.5.2 Regression

Linear regression is used to testify the causal relationship between the dependent and independent variables. Meanwhile, a stepwise method is chosen to find out the comparative importance of the independent variables to the dependent variable.

5.5.2.1 Dependent variable: stakeholder perspective

Fixing the stakeholder perspective as the dependent variable, the stepwise regression model shows that only two independent variables—performance improvement and performance standards-set—come into the model. The multiple-correlation coefficient of the model is .580, while the adjusted determinant coefficient is .330, both are considered to be high. The formula is constructed based on the coefficient presented in the model.

Table 5.20 Regression Model of Stakeholder Perspective

	B	Sig.	R=.580
Constant	2.360	.000	R ² =.336
Performance improvement	.345	.000	Adjusted R ² =.330
Performance standards-set	.250	.002	
Stakeholder=.345×Performance improvement + .250×Performance standards-set+2.360			

5.5.2.2 Dependent variable: internal process

Table 5.21 is the regression model of internal process as the dependent variable and four factors of the performance management. Similarly, the stepwise method again only selects the two same independent variables into the model. Meanwhile, the model shows rather high multiple-correlation coefficient and adjusted determinant coefficient, .653 and .420 respectively. The formula is also calculated and presented.

Table 5.21 Regression Model of Internal Process

	B	Sig.	R=.653
Constant	2.162	.000	R ² =.426
Performance improvement	.401	.000	Adjusted R ² =.420
Performance standards-set	.249	.001	
Internal Process=.401×Performance improvement + .249×Performance standards-set+2.162			

5.5.2.3 Dependent variable: Learning and Growth

When calculating the regression model between learning and growth and other four independent variables, the stepwise method once more put the same variables into the model, excluding the other two factors, performance appraisal and performance report. A high multiple-correlation coefficient (.687) and adjusted determinant coefficient (.467) indicate the high predictability of the model.

Table 5.22 Regression Model of Learning & Growth

	B	Sig.	R=.687
Constant	1.108	.000	R ² =.473
Performance improvement	.562	.000	Adjusted R ² =.467
Performance standards-set	.241	.003	
Learning and Growth =.562×Performance improvement + .241×Performance standards-set+1.108			

5.5.3 Summary

All the four factors of the performance management model have significant correlations with the three factors of government performance. However, while using the regression model, only two factors—performance improvement and performance standards-set—fit the model, reflecting a comparative important causal influence on government performance.

The results of the correlation and regression confirm hypothesis II (significant causal relationships between performance improvement and the internal process) and III (significant causal relationships between performance improvement and the learning

and growth.), partly confirm hypothesis I (significant causal relationships between two dimensions—performance standards-set and performance report—and the stakeholder perspective).

However, one part of hypothesis I—a high correlation between performances report and stakeholder perspective—is not confirmed according to the results. Although the factor of performance report has been adjusted according to the results of the factor analysis, it is still hard to understand why communication with stakeholders has no significance on satisfaction. The feasible explanation is that Chinese citizens pay much more attention to what the governments have done rather than what the government has communicated to them. Facts are stronger than words. On the other hand, it may also reflect two phenomena; one is that the communication between the government and the stakeholders (mainly the customers) is not sufficient that cannot make the stakeholders notice the importance of political participation to them. The other is that the consciousness of proactively communication with governments among citizens is not very strong in China nowadays. In other words, participation in the operation of governments among Chinese citizens is still very low.

5.6 Differences on the ratings between agency levels

5.6.1 Descriptive analysis

In this part, we make a mean comparison between the three different agency levels in the Chinese local governments to find out whether differences do exist. Level 1 refers to municipality government; level 2 corresponds to district/county government, while level 3 means sub-district/town government. The descriptive analysis mainly concentrates on the mean and SD. Table 5.20 shows a general picture of the ratings on the seven factors at different agency levels.

Table 5.23 Descriptive Analysis Between Agency Levels

	Level	N	Mean	Std. Deviation
Performance standards-set	1	96	4.8333	1.00454
	2	60	5.1976	1.02040
	3	40	4.9536	1.05445
	Total	196	4.9694	1.02675
Performance appraisal	1	96	4.5060	1.20595
	2	60	4.8976	1.07874
	3	40	4.6679	1.09582
	Total	196	4.6589	1.15318
Performance report	1	96	4.3993	1.41720
	2	60	4.8556	1.02945
	3	40	4.8833	1.24161
	Total	196	4.6378	1.28934
Performance improvement	1	96	4.1732	1.16407
	2	60	4.6875	.94709
	3	40	4.5156	.98758
	Total	196	4.4005	1.08633
Stakeholder perspective	1	96	4.9813	.91907
	2	60	5.4467	1.03260
	3	40	4.9750	1.05265
	Total	196	5.1224	1.00118
Internal process	1	96	5.1042	.92223
	2	60	5.3000	1.01481
	3	40	5.0950	1.04267
	Total	196	5.1622	.97560
Learning and Growth	1	96	4.5885	1.13843
	2	60	4.9333	1.10072
	3	40	5.0000	1.26085
	Total	196	4.7781	1.16238

5.6.2 Differences between agency levels

ANOVA is introduced in order to find out whether there are differences of the ratings of the seven factors between the three agency levels. Table 5.24 shows that there are significant differences between the three levels on the ratings of three factors: performance report, performance improvement, and stakeholder perspective, with significance level at .039, .011, and .010 respectively.

Table 5.24 ANOVA on Ratings Between Agency Levels

		Sum of Squares	Df	Mean Square	F	Sig.
Performance standards-set	Between Groups	4.912	2	2.456	2.362	.097
	Within Groups	200.659	193	1.040		
	Total	205.571	195			
Performance appraisal	Between Groups	5.668	2	2.834	2.156	.119
	Within Groups	253.649	193	1.314		
	Total	259.317	195			
Performance report	Between Groups	10.717	2	5.358	3.299	.039
	Within Groups	313.453	193	1.624		
	Total	324.170	195			
Performance improvement	Between Groups	10.433	2	5.217	4.583	.011
	Within Groups	219.689	193	1.138		
	Total	230.122	195			
Stakeholder perspective	Between Groups	9.091	2	4.545	4.707	.010
	Within Groups	186.371	193	.966		
	Total	195.461	195			
Internal process	Between Groups	1.643	2	.822	.862	.424
	Within Groups	183.957	193	.953		
	Total	185.601	195			
Learning and Growth	Between Groups	6.865	2	3.432	2.582	.078
	Within Groups	256.606	193	1.330		
	Total	263.471	195			

After having noticed the existence of significant differences, the next step is to find out where these differences exist, between the municipality and district level or between the district/county and sub-district/town level. To solve this technical issue, a statistical method called S-N-K (Student-Newman-Keuls) is adopted. After harmonizing the unequal group size, the calculation reveals that the significant differences only exist concerning the two factors, performance improvement and stakeholder perspective. The results are as follows.

Table 5.25 Student-Newman-Keuls of Performance improvement

Agency Level	N	Subset for alpha = .05	
		1	2
1	96	4.1732	
3	40	4.5156	4.5156
2	60		4.6875
Sig.		.087	.388

Means for groups in homogeneous subsets are displayed.

Uses Harmonic Mean Sample Size = 57.600.

The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table 5.26 Student-Newman-Keuls of Stakeholder Perspective

Agency Level	N	Subset for alpha = .05	
		1	2
3	40	4.9750	
1	96	4.9813	
2	60		5.4467
Sig.		.973	1.000

Means for groups in homogeneous subsets are displayed.

Uses Harmonic Mean Sample Size = 57.600.

The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table 5.25 reflects that performance improvement in the municipality level agencies significantly lags behind that of the district/county and sub-district/town level agencies.

Table 5.26 shows that stakeholder rating on the district/county level agencies is much better than that on the municipality sub-district/town level agencies. Therefore, hypothesis V (stakeholder perspective significantly differentiates between the government agencies at three levels) is confirmed.

Although the S-N-K statistics does not find significant differences with respect to performance report between the three levels due to an unequal data size, we select data and make a mean comparison directly between the municipality level and

district/county level, and the municipality level and sub-district/town level. Both of the differences are significant ($p < .05$). We also make a trend analysis based on the means in the table 5.23. The mean of the three levels on this factor is 4.3993, 4.8556, and 4.8833 respectively. Clearly, the ratings of the sub-district/town level (4.8833) and district/county level (4.8556) are higher than that of the municipality level (4.3993). From this point of view, hypothesis IV (performance report significantly differentiates between the government agencies at three levels) is partly confirmed.

These facts revealed by the statistics are understandable. The long distance from the citizens to the municipality makes the agencies less sensitive to the actual demands and requirements put on them, which in turn make the municipality agencies quite inactive to reforms. On the other hand, district/county and sub-district/town level agencies are much closer to the community, ensuring the proactive adoption of reforms aiming to increase performance and stakeholder satisfaction.

With regard to the differences on the stakeholder perspective, it seems that a closer linkage yield a higher satisfaction rating. Why do the district/county level agencies receive higher ratings than the other two agencies? To some extent, it is related to the Chinese traditional philosophy (a proper distance leads to beauty). However, a closer linkage means that the government is under very close, direct and even personal supervision. It is common sense that it is impossible for the government to make everyone satisfied. A proper distance may lead to beauty (A famous Chinese saying).

The trend analysis notices clues of the differences on performance report between agency levels. Linkage with the citizen is once again a forceful explanation in this regard.

5.6.3 Summary

The mean comparison analysis reveals that two factors are significantly rated between

the three different agency levels; one other factor is differentiated between the agency levels. Possible explanations relate to the linkage or ties with the citizens, which has typical Chinese characteristics in terms of the governmental structure and function.

5.7 Comparison between two nations

Considering only ten Swedish public servants participated in the survey, it is impossible to make a direct comparison with the Chinese data, consisting of 196 respondents. Hence, we selected 10 respondents from the Chinese data at random, five from the municipality level, two from the district/county level, and three from the sub-district/town level. Afterwards, we mixed the two ten-respondents data and formed a new data consisting of an equivalent amount of respondents from the two countries, making it possible to compare means by using nonparametric statistics.

In this new data, the two groups from Sweden and China can both be treated as two independent samples, which means that the Mann-Whitney analysis is suitable in this regard to make mean comparisons.

With respect to the Swedish sample, the factor convergence is based on the results of the factor analysis mentioned above in this chapter.

5.7.1 Mean comparison on the four factors of the performance management

The Mann-Whitney analysis shows that the Swedish respondents mark significantly higher rating on the four factors of the performance management than the Chinese respondents. The hypothesis VI (ratings on the four factors of performance management are significantly different between Sweden and China.) is strongly supported at significant level lower than .011.

The phenomena can be explained by the introduction of the New Public Management

philosophy in the Swedish government at all levels since the 1990s. In the past one-and-half decades, the Swedish local governments have already initiated performance management. Experiences in this regard have been also achieved and cumulated. Compared with Sweden, Chinese local governments are lagging behind in the performance related reform. The performance reform labeled with the New Public Management is on trial in local governments in China. This is the main reason why the ratings on the four factors of the performance management are much lower in the Chinese sample than the Swedish counterparts.

Table 5.27 Test Statistics of the performance management

	Performance Standards-set	Performance Appraisal	Performance Report	Performance improvement
Mann-Whitney U	4.500	11.500	17.500	4.500
Wilcoxon W	59.500	66.500	72.500	59.500
Z	-3.454	-2.919	-2.461	-3.447
Asymp. Sig. (2-tailed)	.001	.004	.014	.001
Exact Sig. [2*(1-tailed Sig.)]	.000	.002	.011	.000

5.7.2 Mean comparison of the three factors of the government performance

It is interesting that only one dimension (Learning and Growth) receives significantly different rating between the two nations, confirming the hypothesis VII (rating on the learning and growth is significantly different between Sweden and China). In recent years, especially after the late 1990s, the Chinese local governments began to adopt some fundamental ideas from the New Public Management movement prevailing in the western nations at that period time. For instance, paying much more attention to and trying whatever means necessary to meet public demands have been the top priorities to in local government agencies, especially in the eastern provinces of China. To reach this target, it is by no means an easy job. One option the Chinese local government employs is to improve efficiency and effectiveness of the government agency. Initiatives, such as the Government Effectiveness Construction Program, have

been launched in recent years aiming to improve the operational efficiency within the government. That can explain why there are no significant differences of the ratings on the stakeholder perspective and internal process between the two nations.

As to the third dimension, it has been pointed out that the Chinese local governments pay much more attention to the short-term and material achievements while inadvertently ignoring the Learning and Growth factor perspective. For instance, although employee training has been a regular task for all agencies, there are many flaws, e.g. lacking of training needs analysis, lacking of career development program. On the contrary, in Sweden, sustainable development is highlighted especially since 2000. Hence, it is understandable that the rating on the Learning and Growth perspective, aiming at the long run development, is much higher in the Swedish sample than in the Chinese counterparts.

Table 5.28 Test Statistics of the government performance

	Stakeholder Perspective	Internal Process	Learning and Growth
Mann-Whitney U	45.000	29.500	17.000
Wilcoxon W	100.000	84.500	72.000
Z	-.381	-1.560	-2.506
Asymp. Sig. (2-tailed)	.703	.119	.012
Exact Sig. [2*(1-tailed Sig.)]	.739	.123	.011

5.8 Summary of the whole analysis

Now we come to the summary of the analysis of the survey. First, we adopted five statistical methods to make a deeper analysis of the data: factor analysis, correlation, regression, mean comparison, and descriptive analysis.

Second, the factor analysis shows a high construct or theory validity of the questionnaire based on the research model combined with two models from the literature review. Moreover, the scale reliability also turns out very high reliability coefficients on all the seven subscales of the questionnaire in terms of Cronbach's

internal consistency coefficient.

Third, the results of the factor analysis reveal that the factor of performance measurement is broadened and renamed as performance appraisal, which combines the concept of the original factor of performance measurement and part concept, mainly focusing on analysis and review, from the original factor of performance report. The reassembly of this new factor is in conformity with the current management reform, labeled as OMA (Object Management Appraisal), in the Chinese local government.

Fourth, after using correlation and regression analysis, hypotheses II, III, and V are perfectly confirmed, while hypothesis I is only partly confirmed, and hypothesis IV is confirmed to some extent. The causal relationship between two factors of the performance management—performance improvement and performance standards-set, and the three performance dimensions are testified. Meanwhile, the comparative importance of the four factors on the performance dimensions is also verified, with the results that performance improvement and performance standards-set are of the most importance.

Fifth, the practice of the performance management between the three government levels differs. Performance improvement and performance report (to some degree) receive significantly different concerns between the three levels. As to the performance dimension, the analytical result also finds that the stakeholder rating differentiates significantly between the three levels. The most forceful explanation is the distance between the agency and the citizens, discussed in detail in this chapter.

Sixth, as to the cross-national comparison, Sweden does a better job than China in terms of the whole procedures of the performance management, confirming hypothesis VI. Moreover, the Swedish local governments pay much more attention to the learning and growth perspective than do its Chinese counterparts, verifying hypothesis VII. With regards to the other factors of government performance, no

significant differences between the two nations exist.

Chapter 6 Conclusions

The conclusion of the whole dissertation is presented. The applicability of the model, methodological criticism, suggestion for future research, and managerial implications are also presented.

6.1 Summary of the dissertation

Accompanying the New Public Management movement throughout the world, performance related reforms have received great attentions since the early 1980s, symbolized by reforms that took place in the UK and the USA. Since the 1990s, all OECD member states, including Sweden, have carried out government performance management reforms initiated by the NPM literature. Similar to some countries in the Asia-Pacific region, which have implemented the reform, China began to initiate a performance management reform in its local governments in recent years, integrating typical Chinese philosophy and means. However, very few empirical researches have been carried out to confirm the validity of performance management, especially in China, a country with very different political system, governmental structure, and culture. It should be a very valuable task to do a research in China, trying to control the level of criterion-related validity of performance management, the current practices of the performance management in the Chinese local government agencies, the differences of these practices between different agency levels. Moreover, very few researches focus on an international comparison. Therefore, to make a direct comparison between Sweden and China in terms of performance is of great value in the academic field.

The literature review on performance management and performance in government agencies is widely dispersed, without any high agreed or famous theories and models.

Many research results are also in question. However, we managed to collect and classify a heap of academic articles. According to the literature review, performance management is a set of procedures aiming to set standard and to improve performance. Two available models are integrated into one model on which our research model is based. Briefly, performance management procedures are divided into four stages, which are named by us in line with their contents as performance standards-set, performance measurement, performance report, and performance improvement. With reference to government performance, a dozen researchers have put forward their own models, among which the balanced scorecard is widely used in the business world and proved to be valuable and applicable also to government agencies. We chose the Balance scorecard as our criteria while excluding the financial perspective.

Simultaneously with the establishment of the theoretical model, seven research hypotheses were formulated related to the causal relationship between the performance management and performance, and differences of the ratings on the performance management and performance between agencies at different levels.

Based on the research model of the performance management and government performance, a 40-items questionnaire was designed in English, consisting of seven sub-scales representing the seven factors of the two sub-models. The Chinese version was translated according to the English version.

The survey was carried out in two ways, one online, and the other by means of paper-and-pen method. During one-and-half weeks, one hundred and ninety-six Chinese respondents from thirty-five agencies at three government levels in Ningbo municipality provided valid information to the survey. However, the response rate is quite hard to calculate precisely. Meanwhile, ten Swedish respondents take part in the survey with only a 5% response rate.

The data was collected and analyzed with SPSS 11.5. Six statistical methods were adopted according to the research purposes. The results of these statistics show that

the construct validity and reliability of the questionnaire are very high. The numbers of the factors are in accordance with the theory. Interestingly, the original factor of performance measurement is broadened combining with measurement and report, which is consistent with the current situation in China. The factor is renamed with performance appraisal.

Correlation and regression analyses reveal that there are significant causal relationships between two factors of the performance management scale (performance improvement and performance standards-set) and the government performance scale. Hypothesis II and III are confirmed, while hypothesis I is only partly confirmed.

Mean comparison analysis, especially S-N-K, shows that performance improvement and stakeholder rating differentiate significantly across the three government agency levels. Another factor of performance report, narrowed to communication with the stakeholders, is differently rated at significant level of .05 while only comparing data from two levels directly. Both the hypothesis IV and V are confirmed.

Nonparametric statistics is employed to make a comparison between Sweden and China, due to small sample in Sweden. The results show that Sweden implements the performance management much better than China, confirming the hypothesis VI. Moreover, it is found that the Swedish local government performs much better than its Chinese counterpart in terms of the learning and growth perspective, testifying the hypothesis VII.

6.2 Applicability of the models

The results of the factor analysis show that the two models, the performance management model and government performance, are consistent with the theoretical structure.

With regards to the performance management model, two original factors—performance measurement and performance report—reassemble, while the

other two factors remain identical with the theory. Therefore, the model is re-established according to the results of the factor analysis.

Table 6.1 A New of Performance Management Model

<p>I: Performance Standards-Set</p> <ul style="list-style-type: none"> -Understanding the purpose of existence -Identify relevant standards -Select indicators -Set goals and targets -Stakeholder expectation consciousness 	<p>II: Performance Appraisal</p> <ul style="list-style-type: none"> -Refine indicators and define measures -Develop data systems -Collect data -Analyze data -Develop a regular report/review cycle
<p>III: Performance Report</p> <ul style="list-style-type: none"> -Feedback/feed forward loops -Communicate expectation to the outside stakeholder 	<p>IV: Performance Improvement</p> <ul style="list-style-type: none"> -Use data for decisions to improvement -Manage changes -Rewards and sanctions -Training needs assessment -Create a learning organization -Suggestions box

According to the results of the factor analysis, this revised model can be widely used in the Chinese local governments at all levels to collect information on the practices of performance management reform at a specific time. It is useful for managers and superior agencies to find out what has been done, and the extent to which an organization has implemented relating with the timetable or original standard with regard to the performance management.

As far as the government performance model is concerned, it verifies the applicability of the balanced scorecard to be used on government agencies, at least in the Chinese local government. This three-dimension model can be adopted to seek information on general and behavioral performance, if the initiator is not concerned with the specific performance, such as output. Even if the initiator aims to collect overall performance information covering a wide range of performance indicators, this model can still be used as a viable, credible and complementary tool.

When it comes to Sweden, it is not yet confirmed whether the revised model is

suitable for the Swedish local government due to the small sample collected in Sweden.

6.3 Methodological criticism

Undoubtedly, this research has its flaws in terms of the methodology.

Firstly, self-assessment questionnaire cannot guarantee the credibility of the information. The performance management and performance are all quite sensitive topics within the government agencies. Self-assessment on respondents' own agency may lead to subjective and self-enhanced results. Moreover, considering the difficulties to find appropriate and sufficient outside stakeholders to rate the stakeholder perspective of the government performance scale, self-assessment is adopted to be a substitute. This arrangement may distort the actual results.

Secondly, the exclusion of the financial perspective and other specific indicators, mainly outputs, make it hard to portray the whole picture of government performance, confining the criterion-related validity to a rather narrower space.

Thirdly, the sample is not sufficient. Chinese respondents from all the three agency levels are not more than one hundred, especially those from the sub-district/town level agencies, making the validity generalization more difficult. As to the Swedish sample, it is rather too small. Although the nonparametric analysis is employed, the results are not so convincing.

Fourthly, the causal relationship analysis is carried out in an ideal situation without any other outside interferential factors that may have an impact on the relationship. In reality, many factors may moderate or mediate the causal relationship, e.g. culture, politics, etc.

6.4 Future research

The models created in this study, the research method used in this study, and the

results yielded in this study can be all regarded as footstones, and pave the way for the future researches. On the other hand, the methodological flaws mentioned above also show implications of adjustments for the researches in the future.

Firstly, more factors besides performance management and government performance should be taken into account while studying the criterion-related validity. For instance, moderators should be introduced into the theoretical frame structure. Some researchers have pointed out the impacts of politics, culture, uncertainty, etc, on causal relationships. The addition of the moderators can make the research results more referential.

Secondly, the government performance dimensions should be broadened. Output and other specific indicators should be put into the performance dimensions, or the criterion-related validity will be confined to a comparatively small area: organizational behavior.

Thirdly, to make the validity analysis much more precise, taxonomy of the government agencies developed by Wilson (Wilson, 1989) should be considered seriously, though it will make the research more complicated.

Fourthly, the usefulness of managerial techniques used in the performance management field, e.g. benchmarking etc, can be the next focal points in future research. The results will be more valuable to the managers who make decisions on introductions of specific methods into the performance management field.

Fifthly, other research techniques such as case study should be highlighted if the researcher wants to get a deeper understanding of the practice of performance management in government agencies.

6.5 Managerial implications to the Chinese local government

The above analysis clearly proves that performance management is an effective

managerial tool to improve government performance. Furthermore, the analysis reveals that the most important steps of the performance management are performance improvement and performance standards-set. However, the descriptive analysis also shows that the rating of the former is the lowest among all the four factors, reflecting the embarrassing situation in the Chinese local governments. The ignorance of this most important factor is a blunder, hampering all the efforts to improve government performance even if the standards and indicators are perfectly defined and selected. The government should re-orient and pay great attentions to performance improvement.

The analysis also shows that government agencies on the municipality level receive the lowest rating of almost all the factors, especially performance improvement, stakeholder satisfaction, and performance report. The long distance between the citizens and the agencies makes these agencies very inactive to the changing environment, which may lead to low accountability. Therefore, the questions how to make the municipality level agencies more proactive, and more effective should be considered seriously. To shorten the distance between the government and its citizens with the community is out of question one optimal way.

As far as government performance is concerned, the rating score on the learning and growth perspective is comparatively low compared with the other two dimensions. Moreover, the score of this factor from the Chinese data is far behind that from the Swedish data. In other words, sustainable development has not yet received sufficient attention. To establish an accountable and sustainable government, this factor should be regarded as the most important one. Improving the competence of employees, building a sustainable culture, and establishing a learning organization all should be emphasized by governors and managers.

The factor of performance report shows that the consciousness of political participation among Chinese citizens is not very high currently. The government should create and provide more chances for citizens' participation in the operation of

government, widely regarded as a fundamental prerequisite to establishing a civil society.

References

Books:

- Hughes, O. E. (2001). *Public Management and Administration: An Introduction (second edition)* Beijing: Chinese People University Press. 1-92. (in Chinese version).
- O'Donovan, K., Walsh, J. S. & Linehan, M. (2000). Enabling performance management through strategic learning: reflection on the use of the balanced scorecard in the Irish public sector. In Montanheiro, L. & Linehen, M. (Eds.), *Public and private sector partnership: the enabling mix* (pp469-478). Sheffield: Sheffield Hallan University.
- Saunders, M., Lewis, P.& Thornhill, A. (2003). *Research Methods for Business Students*. 3 ed. England: Prentice Hall.

Articles:

- Amaratunga, D. & Baldry, D. (2002) Moving from performance measurement to performance management. *Facilities*, Vol. 20, Num 5/6, 217-223.
- Bernstein, D. J. (2001) Local government measurement use to focus on performance and results. *Evaluation and Program Planning*. 24, 95-101.
- Bevan, S & Thompson, M. (1991) Performance management at the crossroads. *Personnel Management*. 23, Nov, 36-39.
- Boland, T. & Fowler A. (2000) A systems perspective of performance management in public sector organizations. *The International Journal of Public Sector Management*, Vol. 13, No. 5, 417-446.
- Brewer, B. & Huque, A. S. (2004) Performance measures and security risk

- management: a Hong Kong example. *International Review of Administrative Sciences*. 70(1), 77-87.
- Brooks, M., Miine, C. & Johansson, K. (2002) Using Stakeholder Research in the Evaluation of Organizational Performance. *Evaluation Journal of Australasia*, Vol.2, No.1, August, 20-26.
- Bryntse, K & Greve, C. (2002) Competitive contracting for public services: a comparison of policies and implementation in Denmark and Sweden. *International Public Management Review*. Vol. 3, Is. 1, 1-20.
- Budding, G. T. (2004) Accountability, environmental uncertainty and government performance: evidence from Dutch municipalities. *Management Accounting Research* 15, 285–304.
- Candrea, P. J. (2004) Analysis of the field of public management: a response to Kelman, Thompson, Jones and Schedler. *International Public Management Review*. Volume 5, Issue 1, 58-68.
- Cavalluzzo, K. S. (2002) Competition, fee-for-service requirements, and government performance: evidence on the Federal Reserve. *Journal of Accounting and Public Policy*. 21, 315-356.
- Cavalluzzo, K. S. & Ittnerb, C. D. (2004) Implementing performance measurement innovations: evidence from government. *Accounting Organizations and Society*. 29, 243-267.
- Chan, Y. C. L. (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*. Vol. 17, No. 3, 204-221.
- Christensen, M. & Yoshimi, H. (2003) Public sector performance reporting: new public management contingency theory insights. *Government Auditing Review*. Vol. 10, 71-83.
- Courty, P. & Marschke, G. (2004) Benchmarking Performance. *Public Finance and Management*. 4(3), 288-316.
- Courty, P. & Marschke, G. (2003) Dynamics of performance-measurement systems. *Oxford Review of Economic Policy*. Vol.19, Is. 2, 268-300.

- Feller, I. (2002) Performance measurement redux. *American Journal of Evaluation*. Vol. 23, No. 4, 435-452.
- Gruening, G. (2001) Origin and theoretical basis of New Public Management. *International Public Management Journal*. 4, 1-25.
- Hacham, H. & Scheinman, R. (2003) The contribution of the accreditation process to public administration: promotion of vital organizational changes. *Accred Qual Assur*. 8, 276-281.
- Holzer, M. & Yang, K. (2004) Performance measurement and improvement: an assessment of the state of the art. *International Review of Administrative Science*. 70(1), 15-31.
- Hoogenboezem, J. A. (2004) Local Government Performance Indicators in Europe: an Exploration. *International Review of Administrative Sciences*. 70, 51 - 64.
- Kates, J., Marconi, K. & Mannle T. E. Jr. (2001) Developing a performance management system for a Federal public health program: the Ryan White CARE ACT Titles I and II. *Evaluation and Program Planning*. 24, 145-155.
- Kaplan, R. S. & Norton, D. P. (1996) Linking the balanced scorecard to strategy. *California Management Review*, Vol. 39, No. 1, 53-79.
- Kelman, S., Thompson, F., Jones, L. R. & Schedler, K. (2003) Dialogue on definition and evolution on the field of public management. *International Public Management Review*. Volume 4, Issue 2, 1-16.
- Kelly J. M. (2005) The dilemma of the unsatisfied customer in a market model of public administration. *Public Administration Review*. January/February, Vol. 65, No. 1, 76-84.
- Kloot, L. & Martin, J. (2000) Strategic performance management: A balanced approach to performance management issues in local government. *Management Accounting Research*. 11, 231-251.
- Landrum, L. B. & Baker, S. L. (2004) Managing complex systems: performance management in public health. *Public Health Management*. 10(1), 13-18.
- Larbi, G. A. (1999) The New Public Management approach and crisis states. *UNRISD Discussion Paper* No.112.

- Liu, Y. L. (2003) Performance indicators of local government—government of the British, the United States and Japan as a reference to our government. *Master degree dissertation of Zhongshan University*. (in Chinese version).
- Martine, L. L. & Singh, K. K. (2004) Using government performance management data to identify new business opportunities: examples from government services outsourcing in the United States. *International Review of Administrative Sciences*. Vol. 70(1), 65-76.
- Mausolff, C. & Spence, J. (2003) Performance measurement and performance effectiveness: a structural equation modeling approach. *Submitted to the International Journal of Public Administration, October*.
- Milakovich, M. E. (2003) Balancing customer service, empowerment and performance with citizenship, responsiveness and political accountability. *International Public Management Review*. Volume 4, Issue 1, 61-83.
- Mol, N. P. & de Kruijf, J. A. M. (2004) Performance Management in Dutch Central Government. *International Review of Administrative Science*. 70, 33-50.
- Murdock, A. (2004) Stakeholder theory, partnership and alliances in the health care sector of the UK and Scotland. *International Public Management Review*. Vol. 5, Is.1, 21-40.
- Mwita, J. I. (2000) Performance management model: a systems-based approach to public service quality. *The International Journal of Public Sector Management*. Vol. 13, No. 1, 19-37.
- O'Toole L. J. Jr. & Meier, K. J. (2003) Public management, personnel stability and organizational performance. *Journal of Public Administration Research and Theory*. 13, 1, 43-64.
- Phillips, J. K. (2004) An Application of the Balanced Scorecard to Public Transit System Performance Assessment. *Transportation Journal*. Vol.43, Is. 1, 26-30.
- Phillips, P. P. & Phillips, J. J. (2004) ROI in the Public Sector: Myths and Realities. *Public Personnel Management*. Vol. 33, Is. 2, 139-150.
- Tat-Kei, A. & Coates, P. (2002) Citizen participation: legitimizing performance measurement as a decision tool. *Government Financial Review*. April, 8-10.

- Theron, C. C. & Spangenberg, H. H. (2000) Confirmatory factor analysis of the performance management audit questionnaire. *South African Journal of Psychology*. Dec2000, Vol. 30 Issue 4, 32-38.
- Thiel, V. S. & Leeuw, F. L. (2002) The performance paradox in the public sector. *Public Performance & management Review*. Vol. 25, No. 3, 267-281.
- Togafau, O. M. (2001), an exploratory study of the American Samoa government's performance measurement and management practices. *A dissertation submitted to Golden Gate University in candidacy for the degree of DPA*. California.
- Wongrassamee, S. Gardiner, P. D. Simmons, J. E. L. (2003) Performance measurement tools: the Balanced Scorecard and the EFQM Excellence Model. *Measuring Business Excellence*, Vol. 7, No.1, 14-19.
- Van Ree, H. J. (2002) The added value of office accommodation to organizational performance. *Work Study*. Vol. 51. Number. 7, 357-363.
- Zeppou, M. & 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*. Vol. 16, Number. 4, 320-332.

Internet:

- Blair, D. (n.d.). Governing for performance in the public sector. Retrieved November 11, 2004, from <http://www.oecd.org/dataoecd/20/1/1813385.pdf>
- Binnendijk, A. (n.d.). Results based management in the development cooperation agencies: a review of experience executive summary. Retrieved November 13, 2004, from <http://ceb.unsystem.org/documents/FB.reports/dacrbm.pdf>
- Bovaird, T. & Halachmi, A. (n.d.) Performance Measurement and Best Value: An International Perspective. Retrieved December 18, 2004, from <http://www.inderscience.com/offer.php?id=94>
- Byrnes, P., Gordon, V. & Jeffers, S. (n.d.) The use of performance measurement tools and techniques by state agencies. Retrieved January 1, 2005, from http://www.abfm.org/pdf_2001_conf/byrnes.pdf
- Chen, G. Q. (2004), (n.d.) The economics foundation of the New Public Management.

- Retrieved March 7, 2005, from <http://www.chinampa.cn/Article/ShowArticle.asp?ArticleID=50> (in Chinese version).
- Fisher, J., Smith, J. & Handley, C. (n.d.) Developing a performance measurement framework for the internal revenue service. Retrieved November 27, 2004, from http://www.calib.com/home/work_samples/files/framework.pdf
- Forbes, M & Lynn, L. E. Jr. (n.d.) How Does Public Management Affect Government Performance? Retrieved December 23, 2004, from http://bush.tamu.edu/content/research/working_papers/llynn/public_manag_affect_gov_perf.pdf
- Guthrie, J. & Steane, P. (n.d.) Stakeholder Analysis of Performance Values: Intellectual Capital in the Third Sector. Retrieved February 26, 2005, from http://www.intcap.com/ICS_Article_2002_Stakeholder_Analysis_of_Perf_Values.pdf
- Moores, G. (n.d.) Framework for assessing performance: outcome-oriented management. Retrieved March 7, 2005, from www.issa.int/pdf/taormina2001/2moores.pdf
- Reichard, C. (n.d.) The Impact of Performance Management on Transparency and Accountability in the Public Sector. Retrieved November 29, 2004, from <http://www.uned.es/113016/docencia/spd%20-%20doctorado%202001-02/Mercado/reichard%20performance%20management%201997.pdf>
- Scotti, S. (n.d.) Moving from performance measurement to performance management. Retrieved December 8, 2004, from <http://www.issa.int/pdf/cracow04/2scotti.pdf>
- Tõnnisson, K. (n.d.) Impact of public sector reform initiatives of performance measurement: the perspective of Estonian local government. Retrieved December 25, 2004, from www.fu.uni-lj.si/egpa2004/html/sg2/Tonnisson.pdf
- Whittaker, J. B. (n.d.) Strategy and performance management in the government. Retrieved November 9, 2004, from www.dmreview.com/whitepaper/WID584.pdf
- Yan, X. R. (n.d.) The New Public Management and its enlightenment on educational reform of primary and middle school in TaiWan. Retrieved March 8, 2005, from_

<http://www.epa.ncnu.edu.tw/epforum/vol11/vol6no1-5.pdf> (in Chinese Version).
Yamamoto, H. (n.d.) New Public Management-Japan's practice. Retrieved November
11, 2004, from <http://www.iips.org/bp293e.pdf>
<http://dict.die.net/government%20agency>, Retrieved March 4, 2005
<http://www.workerscomnsationresourcecenter.com/glossary>, Retrieved March 4, 2005
http://www.investorword.com/3947/public_public_sector.html, Retrieved March 4,
2005

Appendices

Appendix 1

Questionnaire (English version)



Business School, Kristianstad University, Sweden

Government Performance Management Survey

Thank you for your considering to participate in this research survey!

We are international students from Kristianstad University. This survey, a fundamental part of our master degree dissertation, is aiming to learn about the current practices of the performance management in the local government agencies, and the relationship between the performance management and the government performance at agency level.

Your help is critical to our effort to finish the dissertation. If you have any questions or concerns about this survey, please email us at nbsunjun@hotmail.com.

This questionnaire is comprised of 7 headings with a total of 40 questions. All fields are required. Please answer the questions individually. **Please read the interpretations carefully before answering the questions.** It will bother you to take approximately 15 minutes to complete.

Thank you again for your participation.

Personal data

Your country: Sweden China

Age: older than 50 40-50 30-40 20-30

Gender: male female

Working experience (years): 1-5 5-10 10-15 15-20 more than 20

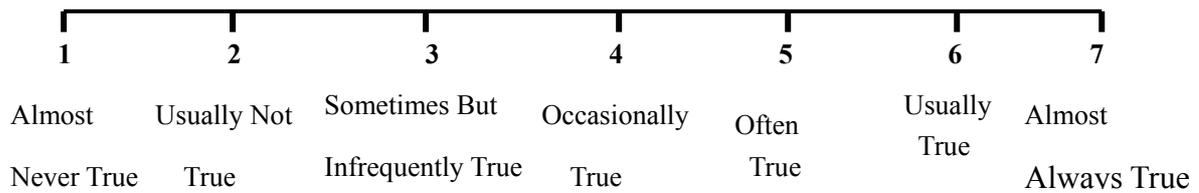
Years in the current agency (years): 1-5 5-10 10-15 15-20 more than 20

Job rank in your agency: top middle low

Education: master or higher bachelor low than bachelor

Size of the agency (people): less than 50 50—100 100—200 200—500

Each heading contains a number of statements, which you should evaluate. A response should be given by ticking the relevant box using the following scoring systems. **Please answer the following questions based on the real situation in your own agency.**



1. Performance management

1.1 Performance Standards-set

This dimension seeks information on the establishment of organizational performance standards, targets, and goals.

Performance standards are guidelines and always the minimum acceptable outcome that are used to assess organizational performance.

Performance indicator is a particular characteristic or dimension used to measure organizational performance.

Performance targets are an intended level of performance expressed as a tangible, measurable objective, against which actual performance can be compared.

No	Statement	1	2	3	4	5	6	7
1	All the employees are familiar with the agency's mission towards the public.							
2	The agency establishes performance standards for every organizational activity.							
3	The agency identifies performance standards through comparing with historical or similar agencies' performance.							
4	The agency selects appropriate performance indicators to assess agency's performance.							
5	The agency sets performance targets for activities to be achieved in a certain time period.							
6	The performance standards, indicators, and targets are communicated within the agency to ensure that all the employees understand.							
7	The agency regularly reports the performance standards, indicators, and targets to the external stakeholders (e.g.: superior authority, taxpayers, consumers, councilors and suppliers).							
8	The agency takes stakeholders (e.g. employees, superior authority, taxpayers, consumers, councilors and suppliers) into account when setting the performance standards.							

1.2 Performance Measurement

This dimension seeks information on the development, application, and use of performance measures to assess achievement of such standards.

Performance measures are quantitative measures of capacities, processes, or outcomes relevant to the assessment of a performance indicator.

No	Statement	1	2	3	4	5	6	7
----	-----------	---	---	---	---	---	---	---

9	Performance indicators are refined to be simple, transparent and precise.							
10	The agency clearly defines ways how to measure organizational activity performance.							
11	The agency establishes a system for collecting and tracking the agency performance data.							
12	The agency measures all or most of the established agency performance standards and targets.							

1.3 Performance Report

This dimension seeks information on the documentation and reporting of progress in meeting standards and targets and sharing of such information through feedback.

No	Statement	1	2	3	4	5	6	7
13	The agency documents the progress related to performance standards and targets.							
14	The agency has a specific system that regularly reports performance data from activities and divisions.							
15	The agency performance data are available to the employees through feedback.							
16	The agency regularly reports the agency's performance information to the external stakeholders (e.g.: superior authority, taxpayers, consumers, councilors and suppliers).							
17	The agency performance data are analyzed and reviewed according to the set performance standards, indicators, and targets.							

1.4 Performance Improvement

This dimension seeks information on the establishment of a program or process to manage change and achieve quality improvement in agency based on performance standards and reports.

No	Statement	1	2	3	4	5	6	7
18	The performance reports are effectively used for decision-making.							
19	The agency performance information is used to set priorities.							
20	The staff has the authority to make certain changes to improve performance.							
21	The agency has specific process to manage changes in policies, programs, or infrastructure.							
22	Employees are encouraged to use different ways to improve the performance.							
23	Rewards and sanctions are related to performance.							
24	The analysis of employee training needs is based on the performance report							
25	The agency has a specific way to collect suggestions on performance improvement from the employees.							

2. Government Performance

2.1 Stakeholder perspective

This perspective seeks information on the overall external stakeholder satisfaction with the goods or services.

External Stakeholder: any outside group or individual who can affect or is affected by the achievement of the agency's objective, e.g. superior authority, taxpayers, customers, councilors and suppliers.

No	Statement	1	2	3	4	5	6	7
26	Customers are satisfied with the responsiveness of the agency in terms of providing goods or services.							
27	Customers are satisfied with the quality of goods or services.							
28	Customers are satisfied with the timeliness of goods and services.							
29	The complaint cases of the stakeholders towards the agency are very rare.							
30	Stakeholders are satisfied with the agency's cooperation with them.							
31	The agency's policy or service has very positive impacts on local social and economic development.							

2.2 Internal Process Perspective

This perspective seeks information on key process excellence.

No	Statement	1	2	3	4	5	6	7
32	The agency clearly knows whom it serves.							
33	The agency can identify the demands of customers in advance.							
34	The agency's service program development is effective.							
35	The service can be efficiently delivered to the customers once the demands are recognized.							
36	The agency's follow-up service effectively guarantees the customers' demands to be fully met.							

2.3 Learning & Growth Perspective

This perspective seeks information on building employees' skills and organizational climate.

No	Statement	1	2	3	4	5	6	7
37	Low-level employees are satisfied with the empowerment.							
38	The employees have the capability to match the demanding environment.							
39	The employees are satisfied with the work climates.							
40	Reliable information is available for strategic decision-making.							

Appendix 2

Questionnaire (Chinese version)



瑞典 Kristianstad 大学商学院

政府绩效管理调查问卷

非常感谢您参与本次问卷调查。

我们(孙骏和徐新桥)正在撰写硕士学位论文,此项调查是论文的重要部分,旨在研究地方政府部门的绩效管理现状,以及绩效管理与政府绩效之间的相互关系。

绩效管理是一个整合了设定绩效标准、选择绩效指标、进行绩效测量和改进提高组织绩效的管理过程。

您的参与将十分有助于我们顺利完成此毕业论文。如果您对本次调查有什么疑问,请通过电子邮件与我们联系 nbsunjun@hotmail.com 和 nbxuxq@yahoo.com.cn。

本问卷共有 2 大部分 40 个问题。

请回答所有的题目。**并请仔细阅读指导语和含义解释。**这可能会打扰您 15 分钟左右的时间来完成本次调查。

再次感谢您的参与!

个人信息:

您的国家: 瑞典 中国

年龄: 50 岁以上 40-50 岁 30-40 岁 20-30 岁

性别: 男性 女性

工作经历: 1—5 年 5-10 年 10-15 年 15-20 年 20 年以上

在本单位的工作年限: 1—5 年 5-10 年 10-15 年 15-20 年 20 年以上

在本单位的职务级别: 高层 中层 一般员工

教育水平: 硕士及以上 本科 大专 中专及以下

单位规模: 50 人以下 50—100 100—200 200—500

指导语:

问卷的每个部分有多个问题, 这些问题需要您根据本单位的实际情况在框中

打“√”，打分请参考以下原则。

1	2	3	4	5	6	7
非常 不符合	中等 不符合	轻微 不符合	不置 可否	轻微 符合	中等 符合	非常 符合

一、绩效管理

(一) 绩效目标设定

本部分旨在了解单位绩效标准和目标等的建立过程。

绩效是指单位的业绩。

绩效标准是指用来测量单位绩效的可以接受的最低水准。

绩效指标是指那些用来测量单位绩效的维度。

绩效目标是指具体的、力求达到的绩效水准，一般要高于绩效标准。

No	问题	非常 不符合		不置 可否		非常 符合		
		1	2	3	4	5	6	7
1	所有员工都非常了解本单位对公众和社会所负有的职责和使命。							
2	单位为每项工作任务都建立了绩效标准。							
3	单位通过与过去绩效或相关单位的绩效作比较来确立本单位的绩效标准。							
4	单位选择了合适的绩效指标来测量本单位的绩效。							
5	单位为工作任务都设定了在一定时间内应达到的绩效目标。							
6	单位告知员工们已经确定了绩效标准、指标和目标等信息，以确保员工们能理解。							
7	单位定期向外部的利益相关者（如上级部门，服务对象，合作对象等）发布已经确定的绩效标准、指标和目标等信息。							
8	单位在设定绩效标准时充分考虑内部员工的特点和公众的期望和要求。							

(二) 绩效测量

本部分旨在了解测量实际业绩过程中测量方式的运用和发展。

绩效测量是指对某个具体绩效指标的量化测量。

No	问题	非常 不符合		不置 可否		非常 符合	
		1	2	3	4	5	6

9	单位采取措施使绩效指标更简化、透明，并能精确反映本单位的绩效。							
10	单位明确规定了测量评估本单位绩效的方法。							
11	单位建立起了收集追踪本单位绩效信息的体系。							
12	单位对已设定的绩效标准和目标都进行测量。							

(三) 绩效报告

本部分旨在了解单位对实际业绩信息的记录和公布，以及反馈和共享这些绩效信息。

No	问题	非常不符合			不置可否			非常符合		
		1	2	3	4	5	6	7		
13	单位根据已设定的绩效标准和目标记录实际业绩。									
14	单位建立了专门制度，定期公布从各部门和各任务活动汇总的绩效信息。									
15	通过单位的反馈，员工们能了解到本单位的绩效信息。									
16	单位定期向外部利益相关者（如上级部门，服务对象，合作对象等）发布本单位的绩效信息。									
17	单位根据已设定的绩效标准、指标和目标对本单位的绩效信息进行分析 and 检查评估。									

(四) 绩效改进

本部分旨在了解单位管理变革与发展的制度，以及如何通过绩效报告提高绩效。

No	问题	非常不符合			不置可否			非常符合		
		1	2	3	4	5	6	7		
18	绩效报告有效地运用在单位各项决策中。									
19	单位的绩效信息被用来设定本单位下一步优先发展的方向。									
20	员工有一定的权力来改变工作方式以提高绩效。									
21	单位建有特定的方法程序去应对在政策、功能或结构等方面的变化。									
22	单位鼓励员工使用不同的方法来提高绩效。									
23	单位的奖惩体系建立在绩效基础之上。									
24	单位根据绩效报告来进行员工的培训需求分析。									
25	单位建有专门机制来收集员工对提高绩效的建议信息。									

二、政府绩效

(一) 利益相关者

本部分旨在了解外部利益相关者对本单位服务的满意度。

外部利益相关者：任何对单位或受本单位影响的外部个人或群体，如上级部门，服务对象，社会公众，合作伙伴等。

No	问题	非常不符合			不置可否			非常符合		
		1	2	3	4	5	6	7		
26	服务对象对本单位在提供服务方面的回应性感到满意。									
27	服务对象对本单位所提供服务的质最感到满意。									
28	服务对象对本单位在提供服务方面的时效性感到满意。									
29	外部利益相关者对本单位的抱怨意见很少。									
30	外部利益相关者对本单位与他们的合作关系感到满意。									
31	本单位的政策或服务对本地区的经济社会发展有非常积极的影响。									

(二) 内部过程

本部分旨在了解单位在内部运作过程中的能力表现。

No	问题	非常不符合			不置可否			非常符合		
		1	2	3	4	5	6	7		
32	单位清楚地知道为谁服务。									
33	单位能预先确认服务对象的需求。									
34	单位服务内容的发展创新是有效的。									
35	一旦掌握了服务对象的需求，单位能高效地把服务提供给这些对象。									
36	单位的跟踪服务有效地确保服务对象的需求得到全部满足。									

(三) 学习与成长

本部分旨在了解单位员工的能力和组织氛围。

No	问题	非常不符合			不置可否			非常符合		
		1	2	3	4	5	6	7		
37	一般员工对单位给予他们的自主权感到满意。									
38	员工们有能力来胜任高要求的工作环境。									
39	员工们对本单位工作气氛感到满意。									
40	单位在进行重大决策时能获得可靠的信息。									

您所在单位的性质： 市级机关 区级机关 乡镇/街道

Appendix 3

Letter to government agency in Kristianstad municipality

Hej,

Jag heter Lisa Nilsson och arbetar på Kristianstad Högskola. Vi har tidigare kommit i kontakt med varandra i samband med att ni var så vänliga och gav en gästföreläsning för våra kinesiska studenter hösten 2004.

Studenterna har haft ett intressant och lärorikt år i Kristianstad, men deras tid i Sverige börjar gå mot sitt slut. De håller nu på att skriva sin uppsats och i samband med sina uppsatser så vill studenterna gärna göra undersökningar i Sverige. De upplever det dock svårt att komma i kontakt med människor och företag som vill ställa upp på deras undersökningar. Två av studenterna, John Sun och David Xu, vill göra en undersökning bland kommunalt anställda tjänstemän och de har därför bett mig att kontakta er. Vi på högskolan skulle uppskatta om ni ville hjälpa dem genom att ställa upp i deras undersökning. Har ni några allmänna frågor så får ni gärna kontakta mig på telefon: 044-20 31 17 alternativt e-post: lisa.nilsson@e.hkr.se

Nedan presenterar studenterna sig och beskriver sin undersökning och dess syfte. Ni finner också en länk till deras frågeformulär. Genom att använda länken och sedan elektroniskt fylla i frågeformuläret så har ni slutfört er del av undersökningen. Ni hittar även kontaktinformation till studenterna för mer specifika frågor.

Deras text och undersökning är av naturliga skäl på engelska. Vi hoppas inte att det ställer till några problem.

Vänliga Hälsningar,

Lisa Nilsson

We are John and David, two Chinese government officials who are studying at Kristianstad University. We are now writing our dissertation on performance management in government agencies. One important part of the dissertation is to collect some empirical data from the Swedish local government departments so that we can compare Sweden and China.

We have designed an online questionnaire which aims to collect information on the performance management practices in your department and the relationship between the performance management and the government performance.

Your participation in our survey is critical to the success of our dissertation. We therefore ask you to please take part in our survey. Considering the requirements of the number of respondents, we also ask you to please send this information to your colleagues and subordinates and ask them to also take part in the survey.

You will find the online questionnaire at the following website:

http://www.dx315.gov.cn/vote/Default_English.htm

Thank you in advance for your participation!

Yours Sincerely,

Sun Jun

Xu Xinqiao

Appendix 4

Letter to government agency in Osby and Hässleholm municipality

Hej,

Jag heter Lisa Nilsson och arbetar på Kristianstad Högskola med ett magisterprogram för kinesiska förvaltningschefer. Programmet är en uppdragsutbildning och inriktat på offentlig förvaltning och internationella affärer. 30 kineser kom till Kristianstad i augusti 2004 och har följt programmet som består av både akademiska kurser samt en mängd studiebesök.

Studenterna har haft ett intressant och lärorikt år i Kristianstad, men deras tid i Sverige börjar gå mot sitt slut. De håller nu på att skriva sin uppsats och i samband med sina uppsatser så vill studenterna gärna göra undersökningar i Sverige. De upplever det dock svårt att komma i kontakt med människor och företag som vill ställa upp på deras undersökningar. Jag hjälper därför studenterna med en inledande kontakt.

Två av studenterna, John Sun och David Xu, vill göra en undersökning bland kommunalt anställda tjänstemän. De ber därför er att hjälpa dem genom att svara på ett frågeformulär. Vi på högskolan uppskattar om ni skulle vilja ställa upp i deras undersökning. Har ni några allmänna frågor så får ni gärna kontakta mig på telefon: 044-20 31 17 alternativt e-post: lisa.nilsson@e.hkr.se

Nedan presenterar studenterna sig och beskriver sin undersökning och dess syfte. Ni finner också en länk till deras frågeformulär. Genom att använda länken och sedan elektroniskt fylla i frågeformuläret så har ni slutfört er del av undersökningen. Ni hittar även kontaktinformation till studenterna för mer specifika frågor.

Deras text och undersökning är av naturliga skäl på engelska. Vi hoppas inte att det ställer till några problem.

Vänliga Hälsningar,

Lisa Nilsson

We are John and David, two Chinese government officials who are studying at Kristianstad University. We are now writing our dissertation on performance management in government agencies. One important part of the dissertation is to collect some empirical data from the Swedish local government departments so that we can compare Sweden and China.

We have designed an online questionnaire which aims to collect information on the performance management practices in your department and the relationship between the performance management and the government performance.

Your participation in our survey is critical to the success of our dissertation. We therefore ask you to please take part in our survey. Considering the requirements of the number of respondents, we also ask you to please send this information to your colleagues or subordinates and ask them to also take part in the survey.

You will find the online questionnaire at the following website:

http://www.dx315.gov.cn/vote/Default_English.htm

Thank you in advance for your participation!

Yours Sincerely,

Sun Jun

Xu Xinqiao