DOCTOR OF PHILOSOPHY

Municipal benchmarking

organisational learning and network performance in the public sector

Marike Noordhoek

2013

Aston University

MUNICIPAL BENCHMARKING: ORGANISATIONAL LEARNING AND NETWORK PERFORMANCE IN THE PUBLIC SECTOR

MARIKE NOORDHOEK

Doctor of Philosophy

ASTON UNIVERSITY

December 2012

Thesis Summary

Aston University

Full title: Municipal benchmarking: organisational learning and network performance in the public sector

Full name: Marike Noordhoek

Degree: PhD

Year of submission: 2012

The main purpose of this dissertation is to assess the relation between municipal benchmarking and organisational learning with a specific emphasis on benchlearning and performance within municipalities and between groups of municipalities in the building and housing sector in the Netherlands. The first and main conclusion is that this relation exists, but that the relative success of different approaches to dimensions of change and organisational learning are a key explanatory factor for differences in the success of benchlearning. Seven other important conclusions could be derived from the empirical research. First, a combination of interpretative approaches at the group level with a mixture of hierarchical and network strategies, positively influences benchlearning. Second, interaction among professionals at the inter-organisational level strengthens benchlearning. Third, stimulating supporting factors can be seen as a more important strategy to strengthen benchlearning than pulling down barriers. Fourth, in order to facilitate benchlearning, intrinsic motivation and communication skills matter, and are supported by a high level of cooperation (i.e., team work), a flat organisational structure and interactions between individuals. Fifth, benchlearning is facilitated by a strategy that is based on a balanced use of episodic (emergent) and systemic (deliberate) forms of power. Sixth, high levels of benchlearning will be facilitated by an analyser or prospector strategic stance. Prospectors and analysers reach a different learning outcome than defenders and reactors. Whereas analysers and prospectors are willing to change policies when it is perceived as necessary, the strategic stances of defenders and reactors result in narrow process improvements (i.e., single-loop learning). Seventh, performance improvement is influenced by functional perceptions towards performance, and these perceptions ultimately influence the elements adopted.

This research shows that efforts aimed at benchlearning and ultimately improved service delivery, should be directed to a multi-level and multi-dimensional approach addressing the context, content and process of dimensions of change and organisational learning.

Key words: Local Government, Municipal Performance Management, Organisational Change, Performance Measurement, Power.

In the absence of my dear mother

1946 - 2007

Acknowledgements

Through my work with local governments worldwide, I have seen various examples of how governments do their best to measure their performance. Despite the good intentions, the use of this information for real performance improvement is not as obvious as the ease with which the idea is embraced. What is it that this noble endeavour in itself is so difficult? This question has encouraged me to explore the science of public administration and governance, looking for more insight. This thesis is an account of this discovery. It was a long journey and not always easy. The journey started with the application process to Aston University when I was living in Salvador de Bahia in Brazil. I am about to finalise my dissertation, but this time living in Maputo, Mozambique with my family. In the intervening years I collected my empirical data in the Netherlands and travelled dozens of times within Asia and Africa for my work as project manager and consultant in the area of local government. And not to forget, the annual trips to Birmingham to discuss progress with my supervisors. The reward for all the hard work however is generous: personal growth and an enhanced understanding and respect for the complexity of organisational learning and network performance in the public sector.

I would like to show my gratitude to the contributions received from my supervisors, Emeritus Professor Stan Brignall and Dr Melina Manochin from Aston University and Dr Sandra Tillema from the University of Groningen. Their research experience formed the foundation on which I could build further. Thank you for all the inspirational support, detailed comments and patience. As Dr Paul Davis was my supervisor from Aston University during the start-up phase of my research study, I want to thank him for his support as well.

I would not have successfully completed my data collection if I was not supported by many people. In particular, my praise goes out to the support from Dr Marcel Hoogwout. I am grateful for his introduction to the building and housing sector and relevant documents shared. I am indebted to the aldermen, managers and employees that I have interviewed in the four case municipalities. They have made this research possible with their openness and honesty about their experiences with municipal benchmarking. I also

want to thank the service providers in the many other municipalities I have met in various countries around the globe, for their role as a source of inspiration and feedback. This also applies to my (former) colleagues at the International Co-operation Agency of the Association of Netherlands Municipalities (VNG International) and at the Quality Institute for Netherlands Municipalities (KING). These colleagues have made it possible for me to combine doctoral research and work for several years. I have always felt that being engaged in both research and practice enabled me to become a better academic. I hope that you as well will reap the benefits of my journey.

I do believe that fun, friendship, solidarity, inner growth and autonomy have helped me in this long journey. Values that some loved ones have supported and encouraged. I want to thank my friends and family for this.

Lukas and Rosa, you were born during the journey that resulted in this thesis. To see you grow up gives me a priceless insight about growth, learning and personal development which cannot be beaten by the writing of a dissertation. I hope that our journey together exposes you to wonder and learn about all features of the elephant (as mentioned in the poem 'the blind men and the elephant'). Finally, I thank Eric, for the attention and the love received during this long journey.

The blind men and the elephant

I.
It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind.

II.
The *First* approached the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl:
'God bless me! - but the Elephant
Is very like a wall!'

III.
The *Second*, feeling of the tusk,
Cried, -'Ho! - what have we here
So very round and smooth and sharp?
To me 't is mighty clear
This wonder of an Elephant
Is very like a spear!'

IV.
The *Third* approached the animal,
And happening to take
The squirming trunks within his hands,
Thus boldly up and spake:
'I see,' quoth he, 'the Elephant
Is very like a snake!'

V.
The *Fourth* reached out his eager hand,
And felt about the knee
'What most this wondrous beast is like
Is mighty plain', quoth he,
'T is clear enough the Elephant
Is very like a tree!'

VI.
The *Fifth*, who chanced to touch the ear,
Said: 'E'en the blindest man
Can tell what this resembles most
Deny the fact who can,
This marvel of an Elephant
Is very like a fan!'

VII.
The *Sixth* no sooner had begun
About the beast to grope,
Then, seizing on the swinging tail
That fell within his scope,
'I see,' quoth he, the Elephant
Is very like a rope!'

VIII.
And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong!

MORAL.
So oft in theological wars,
The disputants, I ween,
Rail on in utter ignorance
Of what each other mean,
And prate about an Elephant
Not one of them has seen!

John Godfrey Saxe (1816 – 1887)

Contents

List of Tables

List of Figures

List of Acronyms

ABC	Aanpak Bouwincidenten Constructieve veiligheid, registration system for construction errors
BWT	Bouw- en Woningtoezicht, Building and Housing Supervision
CBS	Centraal Bureau voor de Statistiek, Statistics Netherlands
CKB	Collectieve Kwaliteitsnormering Bouwvergunningen, Collective Quality Norms for Building Permits
CPB	Centraal Planbureau, Netherlands Bureau for Economic Policy Analysis
EBP	evidence based policy
EBSCO	Electronic Journals Service
EFQM	European Foundation for Quality Management
FTE	full-time equivalent
INK	Instituut Nederlandse Kwaliteit, Dutch Foundation for Quality Management
KING	Kwaliteitsinstituut Nederlandse Gemeenten, Quality Institute for Netherlands Municipalities
KPMG	accounting firm
LTP	Landelijke Toetsmatrix Bouwbesluit, National Test Matrix for the Building Act
Nvivo	qualitative data analysis software
NPM	New Public Management
NVBR	Nederlandse Vereniging voor Brandweerzorg, Dutch Association for Fire and Disaster
NYBMR	New York Bureau of Municipal Research
OECD	Organisation of Economic Co-operation and Development
PM	Performance Measurement
PMI	Public Management Initiative

PMM	Performance Measurement and Management
PMS	Performance Measurement System
PUmA	Programma Uitvoering met Ambitie, Program Execution with Ambition
PwC	PricewaterhouseCoopers
RUD	Regionale Uitvoeringsdiensten, Regional Implementation Services
SGBO	Sociaal Geografisch Bestuurskundig Onderzoek, Social Geography and Public Administration Research
SNN	Samenwerkingsverband Noord-Nederland, Partnership Northern Netherlands
TNS NIPO	market-research firm
TQM	Total Quality Management
VBTB	van Beleidsbegroting tot Beleidsverantwoording, from Policy Budgets to Policy Accountability (performance-based reporting structure)
VNG	Vereniging van Nederlandse Gemeenten, Association of Netherlands Municipalities
VROM	Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer
WABO	Wet Algemene Bepalingen Omgevingsrecht, all-in-one permit for physical aspects
ZENC	Dutch research and consultancy firm

1. Introduction

The research focuses on the relation between municipal benchmarking and organisational learning with a specific emphasis on benchlearning and performance within municipalities and between groups of municipalities in the building and housing sector in the Netherlands. The basic philosophy behind benchlearning is the idea that the foundation of organisational change processes lies in the change of actions and behaviour of individuals and teams (Kyrö, 2003).

Governments in several countries (including emerging economies and developing countries) are demonstrating a growing interest in the area of municipal performance management, and are using performance targets and league tables to obtain indications of the relative performance and efficiency of their organisations. In this context, Dutch municipalities introduced some new public management (NPM) changes in the 1980s and 1990s. Municipal benchmarking is one of the imports from the late 1990s. However, despite the costs associated with it, there is little evidence of results being achieved through the adoption of municipal benchmarking (and related NPM management changes) on organisational learning.

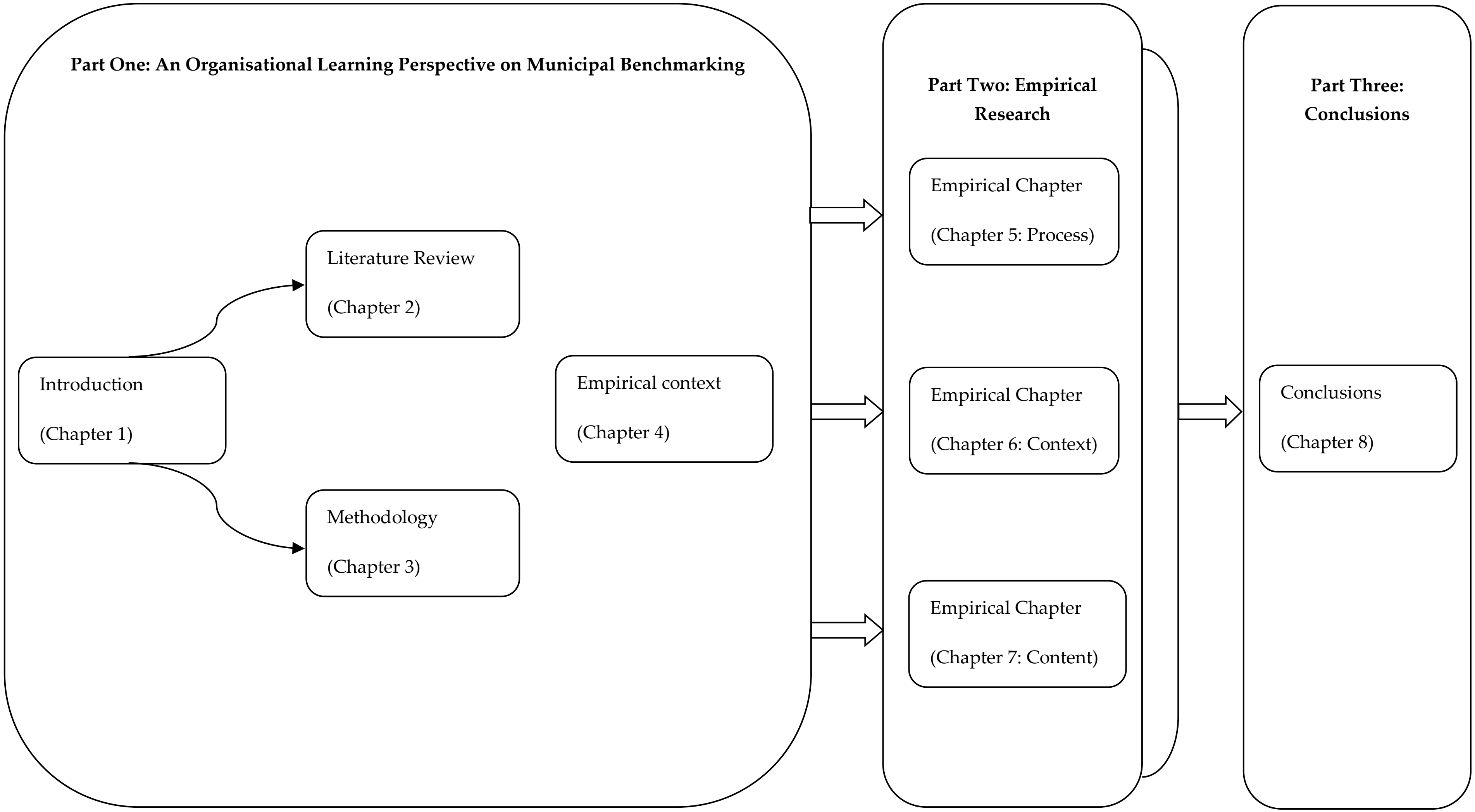
In recent years scholars have investigated topics related to the measurement of performance within the public sector. However, the impact of benchmarking inside the municipal organisation has rarely been studied in an empirical way. Studies on the factors that affect the operation of performance measurement within a local government agency have the character of laundry lists. They do not search for underlying classifications and mechanisms related to the different uses of performance information. Furthermore, in the literature the interpretative theoretical approach did not receive sufficient attention. Studying the "human" side of a municipal benchmarking system allows for further understanding and explaining why people have different experiences with the incorporation and use of performance information. It is suggested in this research study that the relative success of different approaches to dimensions of change and organisational learning are a key explanatory factor for differences in the success of benchlearning.

Two different theoretical approaches have been taken into consideration (organisational change and organisational learning theory) in this thesis, as well as the ontological and epistemological bases of performance measurement and management. The "context, content, process model for successful organisational change" (Pettigrew, 1987) and the "4i framework of organisational learning" (Crossan *et al*., 1999) were used in this research study to assess the underlying classifications and mechanisms to the different uses of performance information.

Moreover, through two distinct phases (pilot interviews and four in-depth cases studies) the relation between municipal benchmarking, organisational learning and network performance is studied empirically. Finally, the classifications and mechanisms related to municipal performance measurement and management are investigated and recommendations are made for both theory and practice.

This introductory chapter provides an overview of the dissertation by presenting the history of performance measurement and management (PMM) in the public sector (1.1) and the rise of benchmarking as a concept (1.2). Section 1.3 provides an introduction to the theoretical background and conceptual frameworks used in this research study. The research focus and the selected areas of interest are presented in section1.4. The last section of this chapter (1.5) gives an insight in the structure of the dissertation. The main phases of the research study and overall structure of the thesis is given in Figure 1.1.

Figure 1.1 Main phases of the research study and structure of the thesis



Source: the author

1.1 Performance measurement and management

Performance ideas have a long history in public management. The early studies on performance measurement in both the private and the public sector were developed by seminal theorists of organisation theory - Winslow Taylor (1911), Max Weber (1922) and Herbert Simon (1946). Taylor studied effective work organisation through time and motion studies, Weber studied the nature of bureaucracy and Simon elaborated on the concept of efficiency, studying performance measures in municipalities (Johnsen, 2005; Shafritz and Hyde, 1992). The first extended implementation of prototypical performance measurement practices arose at the New York Bureau of Municipal Research (NYBMR) after 1906. The roots of the NYBMR practices are primarily the survey, municipal statistics and cost accounting (Williams, 2004). In recent years, the focus of performance measurement in the public sector has narrowed from government activity in general to government services, and its primary purpose has shifted from political accountability to management effectiveness (Williams, 2004, p. 132). Ridgeway elaborated on dysfunctional effects of performance measurement in a well-known paper that appeared in *Administrative Science Quarterly* (1956). Thus, long before the new public management reforms emerged from around 1979 and onwards, dimensions and discourses can be distinguished regarding the implementation, use and outcomes of performance measurement in public management and public sector reforms (Johnsen, 2005, p. 9).

1.1.1 Performance movements

Van Dooren *et al.* (2010) discuss eight movements (i.e., groups of actors sharing a performance agenda) that have propagated PMM. They cluster them into three time segments: (a) pre-World War II, (b) the 1950s-1970s, which roughly parallels the development of the welfare state and the related growth of government, and (c) the 1980s onwards, when welfare states came under pressure from a variety of sources. Table 1.1 gives an overview of performance movements in the twentieth century.

Table 1.1 Performance movements in the twentieth century

Source: Van Dooren *et al*., 2010, p. 45.

With respect to the last time segment it can be said that a clear assumption of NPM was that changes in management systems could and should be made in a way that enhanced performance. According to Moynihan and Pandey (2005) perhaps the most critical single reform of the performance movements was that of performance measurement. The increased data has made it easier to hold managers and management systems accountable for performance. A burgeoning literature also reflects this era of PMM. A common assumption across this literature is that management matters to performance and effectiveness, and that performance is the ultimate goal of public management systems and actions. A widely known literature associated with government reform and NPM was based on these assumptions (e.g., Hood, 1991; Osborne and Gaebler, 1992). Lapsley (2008, p. 93), who examined the future prospects for NPM, concluded that the ultimate determinant of the extent which public administration will achieve, and will succumb to a full blooded NPM, crucially depends on the human frailties of the NPM managers. Modell (2009b), offering a systematic review of institutional research on PMM in the public sector accounting literature, observes that research has moved beyond relatively dated versions of institutional theory, nearly exclusively concerned with isomorphism and symbolic action, by exploring alternative theoretical and analytical paths enriching this perspective. According to the author, this signifies a shift in analytical foci from a one-sided emphasis on the institutional effects on PMM, treating institutional pressures as largely exogenous, to examine the more intricate roles of PMM as an outcome of as well as a medium for change.

1.1.2 Use of performance information

Several authors discussed the uses of performance information as they can be found in organisational practice (e.g., Hatry, 1999; Poister, 2003; Behn, 2003; Greiling, 2005). Greiling (2005) indicates that uncertainties persist as to whether performance measurement should be used internally as an internal management instrument or as an organisational learning instrument. She observes that to mingle those is dysfunctional. According to Greiling (2005, p. 565):

> "*If performance measurement is applied as a monitoring instrument, the units monitored will be anxious to paint a positive picture of their performance; if it is combined with a performance based inducement system, the problem will become even bigger. Organisational learning on the other side, which has been preferred by public bodies in Denmark, the Netherlands and Sweden from the very beginning, makes it necessary that an unveiled picture of the areas for improvement is shown*".

Rantanen *et al*. (2007) observe two additional problems for the performance measurement system in the public sector: (1) taking into account all stakeholders may result in producing a multitude of performance measures that satisfy no one, and (2) it may be difficult to set targets or to make decisions based on the measurement results, because some of the stakeholders have conflicting objectives.

In order to study the use of performance information in organisations, scholars have developed classifications of uses into broader categories with similar features. Behn (2003), for example, proposes a categorisation of eight managerial uses. Van Dooren (2006) classifies the uses in a more limited set of three categories (based on a review of ten texts on performance measurement that yielded 44 uses of performance information): (1) use for research and learning; (2) use for internal management; and (3) use for accountability purposes. Behn (2003, p. 586) observes that public managers should think seriously about the managerial purposes to which performance measurement might contribute and how they might deploy these measures. Only then they can select measures with the characteristics necessary to help achieve each purpose. Van Dooren (2006) adds that the diverse uses of performance information require an adapted design of the measurement system. The classification of performance management and the accompanying features as referred to by Behn (2003), Van Dooren (2006) and Van Dooren *et al*. (2010) are summarised in table 1.2.

Table 1.2 Three clusters of performance management

	Research and learning	Internal management	Accountability purposes
Why	Evaluate Learn Improve	Control Budget Motivate	Promote Celebrate
Key question	How to improve policy or management?	How to be in control of activities?	How to communicate performance?
Focus	Internal	Internal	External
Orientation	Change / Future	Control / Present	Survival / Past

Sources: Behn, 2003; Van Dooren, 2006; Van Dooren *et al*., 2010.

A crucial decision for defining the focus and orientation of performance management is whether performance information will be used in a hard or a soft way (De Bruijn, 2007; Van Dooren *et al*., 2010). Van Dooren *et al*. (2010) refer to two dimensions of this distinction:

1. Coupling between performance information and judgement: hard use presupposes a tight coupling between performance information and judgement, while soft use leaves more room (the difference between formula-based use and interpretative use);
2. Consequences of the judgements that are based on the performance information (low or high impact).

Brignall and Modell (2000) suggest that decoupling of (as opposed to integration between) performance indicators and goals is a viable strategy for seeking simultaneous legitimacy of multiple constituencies. Johnsen (1999) concluded from a case study of four Norwegian municipalities that in complex settings a decoupled information mode may be more successful than a coupled one.

Another observation that can be made in relation to the use of performance information is the non-use of performance information. Performance information is often not picked up despite its potential benefits. Van Dooren *et al*. (2010) give four explanations why performance information is not used on those instances where it is available and even incorporated. The first explanation of non-use is about the technical quality of the information itself. It is suggested here that decision-makers when confronted with new information are performing two tests: a truth test and a utility test. Decision-makers appraise the truth of information in terms of its technical merit and by checking the conformity of the findings with their prior understanding and experience. Utility of the information is assessed by the extent to which it provides explicit and practical direction and challenges current practices on matters decision-makers can do something about. The second explanation about non-use is about the user of the information. Based on the theory of Herbert Simon (the model of the Economic man) the authors suggest that performance measurement does have the potential to improve decision-making. However, this potential will only materialise when performance measurement professionals explicitly acknowledge the existence of bounded rationality, and do not take rationality for granted. The observation that performance measurement will operate in a fundamentally different way according to the culture is a third explanation of non-use. The causes of the non-use are sought in the mismatch between the use of performance information and the cultural traditions of a society, administration or organisation. The authors refer to a study of Sartori (1969) describing the traditional distinction between cultures: learning from performance information may be easier in pragmatic, empirical cultures than in rationalist, deductive ones. The last explanation given by Van Dooren *et al*. is about the influence of the institutional context on non-use of performance information. A distinction is being made between regulatory institutions (reflecting the power distribution in a politico-administrative arena) and normative institutions (the values, norms and roles that guide behaviour).

1.1.3 Effects of performance measurement

The positive effects of performance measurement are mentioned in several publications (e.g., De Bruijn, 2007; Holzer and Yang, 2004; Holzer and Kloby, 2005). It can be summarised that performance measurement promotes transparency and innovation, is an incentive to be productive, may help to de-bureaucratise an organisation, promotes learning and may enhance an organisation's intelligence position. There is another picture, however, apart from this beneficial effect of performance measurement, which is that performance measurement creates a large number of perverse effects. Van Dooren (2006) observes that use for accountability purposes will lead to the highest perverse effects as opposed to using the performance information for management or learning. Several authors (Fryer *et al*., 2009; De Bruijn, 2006; Smith, 1995; Adcroft and Willis, 2005; Julnes and Holzer, 2001; Pidd, 2005; Van Thiel and Leeuw, 2002) list types of deviant behaviour and their causes. An overview of types of deviant behaviour and their causes is given by Fryer *et al*. (2009) in the table below.

Table 1.3 Types of deviant behaviour and their causes

Source: Fryer *et al*., 2009, p. 486.

1.1.4 Defining performance measurement and management

Performance measurement and performance management have accumulated many definitions. Greiling (2005) discerns three common definitions in the performance measurement literature. In the first definition, performance measurement is limited to applying various techniques for generating performance data. Second, performance measurement is used in the sense that it refers to performance reporting. This function includes the use of performance measurement as an accountability tool. Third, performance measurement is advocated as a steering instrument within the public sector. Some authors refer to the last two functions of performance measurement by the term performance management. Van Dooren (2006) and Van Dooren *et al.* (2010) make a clear distinction between measurement, incorporation and use of performance information. Van Dooren's classification divides performance management as referred to by Greiling in her second and third definition into two steps. A systematic overview is given in table 1.4 below.

Table 1.4 Performance measurements versus performance management

Performance measurement	Measuring	Systematically collecting data by observing and registering performance-related issues for some performance purpose.
Performance management	Incorporating	Intentionally importing performance-related data in documents and procedures with the potential and purpose of using them.
	Use	Debates and institutionalised procedures for the purpose of designing policies, for deciding, for allocating resources, competencies, and responsibilities, for controlling and redirecting implementation, for (self) evaluating and assessing behaviour and results, and for substantiating reporting and accountability mechanisms.

Source: the author (based on Crossan *et al.*, 1999; Greiling, 2005; Van Dooren, 2006 and Van Dooren *et al.*, 2010)

The classification of definitions as presented in table 1.4 will be used in the remainder of this thesis, allowing for a clear understanding of the difference between performance measurement and performance management. The subsequent section gives insight in the historical development of benchmarking and into benchmarking typologies according to their function.

1.2 Benchmarking

1.2.1 Historical development of benchmarking

An understanding of the historical development of benchmarking is important to see if benchmarking means the same today as it did in the past and to identify the scope of benchmarking. Benchmarking was developed as a private sector instrument under the umbrella of Total Quality Management (TQM). The term *benchmarking* was coined in Xerox. The approach was highly successful at this company, and by 1981, benchmarking was adopted as a corporate-wide effort. Benchmarking at that time was built on competitive analysis (Andersen *et al*., 2008).

The generations of benchmarking in the period between 1940 and 2010 reflect seven approaches to benchmarking (see figure 1.2 below). Kyrö (2003) gives a brief summary of the first five generations of benchmarking in the private sector that were initially classified by Watson (1993):

> *"The first one, entitled "reverse engineering", was product oriented, comparing product characteristics, functionality, and performance of competitive offerings. Second generation "competitive benchmarking" involved comparisons of processes with those of competitors. Third, process benchmarking was based on the idea that learning can be made from companies outside their industry. Hence sharing of information became less restricted, non-competitive nature of intelligence gathering. But at the same time it required more in-depth understanding and needed to understand similarities in processes, which on the surface appear different. Fourth generation, in the 1990s, introduced strategic benchmarking, involves a systematic process for evaluating options, implementing strategies and improving performance by understanding and adopting successful strategies from external partners. Typical to this perspective is continuous and long-term development and the aim to make fundamental shifts in process. With fifth generation this was complemented by global orientation (Kyrö, 2003, p. 213)".*

The most recent studies of benchmarking (Moriarty and Smallman, 2009; Andersen *et al*., 2008; Anand and Kodali, 2008; Kyrö, 2003) provide two new approaches: benchlearning (or competence benchmarking) and network benchmarking. According to Kyrö (2003) the basic philosophy behind benchlearning is the idea that the foundation of organisational change processes lies in the change of actions and behaviour of individuals and teams. In her study Kyrö refers to a publication by Karlöf and Östblom (1995) in the Finnish language in which the authors use the term benchlearning. Benchlearning refers to a cultural change in efforts to becoming a learning organisation. Organisations can improve their effectiveness by developing competences and skills by learning how to change attitudes and practices. Network benchmarking is seen by several authors (e.g., Moriarty and Smallman, 2009; Kyrö, 2003) as a new type of benchmarking. Instead of one single unit or organisation, benchmarking might involve a network on both sides.

Figure 1.2 Generations of benchmarking

Aston University

Illustration removed for copyright restrictions

Source: Kyrö, 2003, p. 214.

In the late 1980s and early 1990s benchmarking made its appearance in the public management literature as a method to improve performance and accountability. In the subsequent section an introduction is given to public sector benchmarking.

1.2.2 Public sector benchmarking

The book *Reinventing Government* (Osborne and Gaebler, 1992) initiated the enthusiasm for public sector benchmarking in the United States. In Europe the Organisation of Economic Co-operation and Development (OECD) hosted an international meeting dedicated to public sector benchmarking in 1997 (Braadbaart and Yusnandarshah, 2008).

Improved performance was defined by the private sector as reduced error or costs, increased profit margins, or increased market share. Kouzmin *et al*. (1999, p. 122) observe that the motivational forces and obstacles are somewhat different in the public sector. Benchmarking is supposed to introduce competition into a state apparatus context that is characterised by the cooperation of public sector agencies for the "collective" public good. Public services operate with a fixed budget and consumer groups are in competition with each other for scarce resources. In this situation, consumer satisfaction cannot be the only, or dominating, dimension in performance measurement in the public sector. Whereas in the 1980s the focus was on the "three Es", economy, efficiency and effectiveness, in the 1990s attention has shifted to quality and consumer satisfaction.

Tillema (2007) argues that important characteristics of public sector organisations may influence these organisations' use of benchmarking information for performance improvement. The characteristics the author refers to are: limited market exposure; institutional constraints; formal and informal influences; public scrutiny; complex objectives and barriers to innovations. Based on four explorative case studies of Dutch water boards the author concludes that these characteristics of public sector organisations can be expected to affect performance improvement negatively.

1.2.3 Defining benchmarking

In the literature on benchmarking several definitions are given to describe the concept (e.g., Kouzmin *et al*., 1999; Bowerman *et al*., 2002; Folz, 2004; Van Helden and Tillema, 2005; Ammons *et al*., 2001; Wynn-Williams, 2005; Ammons and Rivenbark, 2008;

Goncharuk and Monat, 2009). All definitions imply that benchmarking is a process and includes measurement, comparison, implementation and improvement.

Bowerman *et al*. (2002) present two public sector benchmarking typologies in an attempt to distinguish emerging practices in the public sector from practices associated with practices in the private sector. The authors use the term "compulsory benchmarking" to reflect the use of benchmarking data by other agencies to comment publicly on the performance of public sector bodies. "Voluntary benchmarking", conversely, indicates the use of benchmarking by public sector managers as an approach to improving performance *per se*. At the end of the same decade Braadbaart and Yusnandarshah (2008) studied the public benchmarking literature between 1990 and 2005. According to the authors a theoretical and conceptual rift runs through the literature, with those advocating public sector benchmarking as a tool for managed competition on one side, and those promoting benchmarking as a voluntary and collaborative learning process on the other. In several articles it is concluded that as long as benchmarking is conflated with this external monitoring function, and used for managed competition, its potential to promote a voluntary and collaborative learning process on the other hand will be difficult to achieve (Ball *et al*., 2000; Bowerman *et al*., 2002; Northcott and Llewellyn, 2005; Papaioannou, 2007).

The above discussion raises a fundamental question: what is benchmarking for? Section 1.1.2 introduced the uses of performance information in three categories: (1) use for research and learning; (2) use for internal management; and (3) use for accountability purposes. This research study focuses on the first two categories (voluntary benchmarking). Plotting the two public sector benchmarking typologies according to their uses, the following classification can be made:

Table 1.5 Benchmarking typologies according to their uses

Benchmarking typologies	**Voluntary benchmarking**		**Compulsory benchmarking**
Use	Research and learning	Internal management	Accountability purposes
Why	Evaluate Learn Improve	Control Budget Motivate	Promote Celebrate
Key question	How to improve policy or management?	How to be in control of activities?	How to communicate performance?
Focus	Internal	Internal	External
Orientation	Change / Future	Control / Present	Survival / Past

Source: author (adapted and modified from Behn, 2003; Van Dooren, 2006; Van Dooren *et al*., 2010)

1.3 Theoretical approach

An extensive review of the literature (see chapter 2) showed that the main academic contributions in the area of performance measurement, performance management and benchmarking have a positivist theoretical orientation. Two key observations can be made. The first observation is that the effective use of performance information derived from municipal benchmarking requires a major change in the attitude of the people using the information. According to Coplin *et al*. (2002), a game of incessant organisational change is the real danger of benchmarking and not the perverse effects. Studies in this field indicate that most benchmarking efforts are hampered by resistance from both managers and lower level employees to change. The second observation is that the studies of the learning potential of a benchmark may be strengthened when insight is given in how a municipal benchmark shapes organisational learning within local government agencies.

Bourne *et al*. (2003b, p. 263) suggest that successful implementations of performance measurement systems should draw from change management insights and should not be confined to the limited prescriptions offered by the performance measurement literature. The authors observe that the further the change proceeds, the less informative the performance measurement literature becomes. In addition to this, Dutta and Crossan (2003) note several parallels between change and organisational learning. For example the observation that the context, content and process of change correspond to respectively situated, cognitive and behavioural dimensions of learning. The authors developed a theoretical perspective on how organisational learning informs and extends the understanding of change. It is supposed that in order to accomplish change, organisations must necessarily learn. Dutta and Crossan (2003) suggest that there is potential for significant cross-fertilisation of ideas from these schools. As a consequence, this thesis draws on a conceptual framework linking organisational change (1.3.1) and organisational learning theory (1.3.2).

1.3.1 Pettigrew's model for organisational change

In understanding and evaluating the impact of municipal benchmarking, it is important to understand where and how impact is created. This means that it is important to interpret and model the context within which impacts occur. Several authors see change either as a single event or as a discrete series of episodes that can be de-contextualised. Burnes (2000) argues that the planned organisational development approach that derives from Kurt Lewin's model of change (unfreezing, changing and refreezing), dominated thinking from the 1940s to the early 1980s. He claims that since the 1980s there has been increasing criticism of this approach, especially in the more contextual and processually oriented studies. For Pettigrew (1985), change and continuity, process and structure, are inextricable linked. Pettigrew is highly critical of the organisational development approach to change and is seen to ignore the importance of changing. The processual perspective claims that an understanding of power and politics is central to an understanding of the processes of organisational change. The foundational work of Pettigrew (1985) has been widely referenced and discussed in the organisational change literature (e.g. Buchanan and Storey, 1997; Burnes, 2000; Dawson, 1994; Kanter *et al*., 1992).

In 1987, Pettigrew presented the "context, content, process model for successful organisational change". The author explains that the 'how' of change can be understood from an analysis of process, while much of the 'why' of change is derived from an analysis of inner and outer context and that the 'what' of change is encapsulated under the label content.

Inquiry into the link between strategic decision making (choice processes) and performance traditionally has been divided into process (how) and content (what) research. Pettigrew (1987) suggests that the role of the context (both internal and external to the organisation) in which process and content are embedded, must be considered. The result is a framework that incorporates three time frames, namely past, present and future time. Time is thus deemed as vital to the unfurling of a particular process (Hinings, 1997, pp. 498-500).

The <u>process</u> of change refers to (Pettigrew, 1987, p. 658):

> *"The process of change refers to the actions, reactions, and interactions from the various interested parties as they seek to move the organisation from its present to its future state".*

Bourne *et al*. (2003a) reviewed the change management literature to create a better understanding of why so many attempts to implement a PMS fail. Their review suggests that one difficulty might lie in the fact that the PMS design processes described in the literature are only partial processes. They create the desire for change and provide the first steps for change, but give little guidance on implementation. Bourne *et al*. (2003b, p. 261) conclude that a successful implementation of PM systems requires both hard (rational objectives) and soft systems (human objectives) of change, but with a stronger soft systems component. Dutta and Crossan (2003) observe that the process school is focused on studying change as a phenomenon that transforms the organisation. The authors refer to a study of Carroll and Hannan (2001, p. 358), who say: "*content change refers to what actually differs in the organisation at the two points in time whereas process change is the way the change in content occurs, the speed, sequence of activities, decision-making and communication systems deployed, and the resistance encountered*". Longbottom (2000) reviewed over 460 articles on benchmarking. From his analysis, he concludes that benchmarking as a process is similar to the quality cycle presented by Deming involving a continuous process of plan, do, check and act.

The context of change refers to (Pettigrew, 1987, pp. 655-657):

> "*A contextual analysis of a process draws on phenomena at vertical and horizontal levels of analysis and the interconnections between those levels through time. The vertical level refers to the interdependencies between higher or lower levels of analysis upon phenomena to be explained at some further level; for example, the impact of a changing socioeconomic context on features of intra-organisational context and interest-group behaviour. The horizontal level refers to the sequential interconnectedness among phenomena in historical, present, and future time. (…) Outer context refers to the social, economic, political and competitive environment in which the organisation operates. Inner context refers to the structure, culture, and political context within the organisation through which ideas for change have to proceed.*"

Moynihan and Pandey (2005) seek to conceptualise and empirically test how inner and outer contextual changes combine to create performance. They conclude that the following external environmental variables have a positive impact on performance: elected officials' support; public and media influence; gubernatorial power and the availability of resources. In addition, they conclude that the following internal management choices are positively associated with performance: ability to create a developmental organisational structure; establish a focus on results through goal clarity and customer-service training; decentralizing decision-making authority and improvements in information technology capacity and job satisfaction.

With reference to organisational learning theory, Rashman *et al*. (2009) argue that particular combinations of external and internal contextual factors may lead to sector-specific learning drivers, goals, needs, structures, systems, practices and outcomes. According to the authors it is important to describe the context-specific factors for a level of learning and to describe the nature of different participating organisations, as well as the network structure itself.

The content of change refers to (Pettigrew, 1987, p. 657):

"Content refers to the particular areas of transformation under examination. Thus the organisation may be seeking to change technology, manpower, products, geographical positioning, or culture".

From an organisational learning perspective Dutta and Crossan (2003, p. 6) add that:

"C*hange as content looks at organisational parameters before and after an event in the life of the organisation and tries to understand what the possible antecedents and consequences of the change could have been. In other words, change as content assumes that change is a tractable phenomenon and that it is almost possible to halt the organisation in its track as it were, in order to define and engage in the collection of relevant measurements that help us to study change*".

So, content is concerned with the areas of transformation and the tools and techniques used to effect change. A performance measurement system (including the parameters of the system itself) can be seen as such a tool, possibly effecting changes in organisational characteristics including performance. In this research study, a performance measurement system is seen as a possible antecedent of the change. This research study focuses on one aspect of content, i.e. one of its antecedents. The content of benchlearning in this thesis refers to the influence of different definitions of performance used by various institutional actors, allowing for a focus on the quality of achievement and/or the quality of actions. In chapter 2, the literature review, the broader perspective of Pettigrew is used.

The use of the framework of Pettigrew for successful organisational change has three main advantages. Firstly, the framework explicitly addresses context and action as a guiding assumption. Secondly, the framework explicitly addresses a search for holistic rather than linear explanations of processes. Thirdly, the framework explicitly addresses embeddedness and temporal interconnectedness by studying processes across a number of levels of analysis and studying processes in past, present and future time (Pettigrew, 1997, p. 340).

1.3.2 4i framework of organisational learning

In understanding and evaluating the impact of municipal benchmarking, it is important to understand the relation between dimensions of change and approaches of learning. It is supposed that in order to accomplish change, organisations must necessarily learn. This means that it is important to interpret and model the learning processes within which impacts occur. An appealing model that can be used for this purpose is the "4i framework of organisational learning" presented by Crossan *et al*. (1999).

In their framework learning processes of intuition-interpretation-integrating-institutionalisation interact with knowledge exploitation and exploration dynamics throughout three levels of analysis: individual, group and organisational levels.

Dutta and Crossan (2003) have synthesised the elements as identified in the 4i framework to develop a dynamic, integrative model of organisational change, including the different levels and social and psychological processes of learning. The authors relate the dimensions of change as formulated by Pettigrew (process, context, content) to different approaches of learning (respectively behavioural, situated and cognitive learning). See table 1.6 below for an overview.

Table 1.6 Relation between dimensions of change and approaches of learning

	How?	Why?	What?
Dimension of change	Process	Context	Content
Approach to learning	Behavioural	Situated	Cognitive

Source: author

The authors give a description of the different approaches to learning in their article. The *behavioural* side of learning involves construction of identity and the cultural and interpretive dimensions of organisational learning. Learning as *situated* practice recognises that learning occurs in individuals and teams when a group of people engage in a series of activities when placed in a particular situation or context. The *cognitive* perspective regards the treatment of organisational learning as the summation of learning of individuals (Dutta and Crossan, 2003, pp. 3-8).

Four key premises form the underpinnings of this framework and support one central proposition:

- Organisational learning involves a tension between assimilating new learning (i.e., exploration) and using what has been learned (i.e., exploitation).
- Organisational learning is multi-level: individual, group, and organisation.
- The three levels of organisational learning are linked by social and psychological processes: intuiting, interpreting, integrating, and institutionalising (4i's).
- Cognition affects action (and vice versa).

The central proposition of their framework is that the 4i's are related in feed-forward and feedback processes across the levels.
Intuition is a uniquely individual process. It may happen within a group or organisational context, but the recognition of a pattern or possibility comes from within an individual. Organisations do not intuit. *Interpretation* is a personal explanation for one's self and for others about individual insights. This is a learning process that requires verbal manifestations and language development. *Integration* is a process through which shared understandings among individuals occur and coordinated actions through mutual adjustments result. For this to happen, dialogue and conjoint action are crucial. Initially this will be an informal and ad hoc process but if coordinated action is recursive and meaningful and if managers recognise it as a valuable practice, it might be institutionalised (Crossan *et al*., 1999, p. 525). *Institutionalisation* is a process of guaranteeing that routinised actions occur. For this to happen, tasks, beliefs, values, norms, roles and responsibilities are defined, actions specified, and organisational mechanisms put in place. Institutionalisation is the process of embedding individual and group learning within the organisation including its symbolic and relational systems, structures, procedures and strategies (Crossan *et al*., 1999; Wiseman 2008).

The four learning processes occur over the three learning levels. Although the definitions are presented in a linear fashion, it is critical to appreciate the iterative nature of the processes. Intuiting occurs at the individual level and institutionalising at the organisational level; however, interpreting bridges the individual and group levels, while integrating links the group and organisational levels (Crossan *et al*., 1999, p. 525).

The use of the 4i framework is further justified in chapter 5 where reference is made to different theoretical conceptions and other frameworks that are available. The use of the 4i framework of organisational learning has two main advantages. Firstly, the framework explicitly relates dimensions of change to different approaches of learning. Secondly, the framework explicitly addresses the importance of a multi-level approach by studying approaches of learning at the individual, group and organisational levels.

Figure 1.3 below summarises the 4i framework of organisational learning.

Figure 1.3 The 4i framework of organisational learning

Aston University

Illustration removed for copyright restrictions

Source: Crossan *et al*, 1999.

Lawrence *et al*. (2005) integrated power into the 4i framework of organisational learning. The authors argue that power and politics provide the social energy that transforms the insights of individuals and groups into the institutions of an organisation. According to them, the political will and skill of those attempting to make a transformation of new ideas into coherent collective action is a key issue in understanding which ideas will be integrated into the activities of groups and which will become institutionalised in organisations. A slightly modified version of their framework is well suited for the analysis of the empirical data presented in chapter 6, because it relates forms of power in organisations to specific learning processes. In addition to the conceptualisation and typology of power as formulated by Lawrence et al., chapter 6 relies on the typological classification of Miles *et al*. (1978), used to analyse forms of strategic stances. Their classification suggests that to solve their problems, organisations employ four strategic types: defenders, analysers, prospectors and reactors. A further explanation of these classifications is offered in chapter 6.

1.4 Research focus

The review of the literature and the research gaps identified (chapter 2) enabled the identification of a number of opportunities for further research (see section 2.2.3). The linkages between the identified areas, the relevance of the topics, and the researcher's personal interests are the main reasons for determining the areas of particular focus for this research. These areas are seen as crucially important for the further development of PMM and benchmarking in the public sector. This research focuses particularly on the three following issues:

1. Process: the interconnections between organisational learning and the incorporation and use of benchmark information (chapter 5);

2. Context: the influence of different forms of strategic stances and power on the learning outcome of benchmarking and improved performance (chapter 6);

3. Content: the influence of different definitions of performance by various institutional actors on the use of performance information from benchmarking (chapter 7).

The first area of interest (process) suggests that studies into the learning potential of a benchmark may be strengthened when insight is given in how a municipal benchmark shapes organisational learning within local government agencies. The literature suggests that municipal benchmarking is likely to deliver improved organisational performance when it is coupled to organisational learning (Argyris and Schon, 1978; Moynihan, 2005; Kouzmin *et al*., 1999; Rashman *et al*., 2009).

The second area of interest relates to context. The influence of different forms of strategic stances and power is crucial for organisations that depend on, and have to report to and engage with, a number of different parties. This is especially the case with public sector organisations, where others "exercise an influence upon the way the rules of the game are formulated, and how it is played out in the public domain" (Boivard and Loeffler, 2003, p. 167). Empirical research is required to understand how different forms of strategic stances and power influence benchlearning.

The third area of interest (content) is particularly relevant in a public sector context, since, despite the increasing use of PMM and benchmarking, empirical research is required to understand how performance is defined by different stakeholders, addressing the theoretical and conceptual rift running through the literature (i.e., benchmarking as a vehicle for competition or cooperation) and its influence on benchlearning (Behn, 2003; Braadbaart, 2007; Van Dooren, 2006).

1.5 Structure of the text

In the next chapter, chapter 2, the literature is reviewed in order to position this research study vis-à-vis recent studies. Chapter 3 presents the empirical research questions (3.2) and offers insight in the philosophical perspective and methodology used in this research study. In the fourth chapter, the empirical context is described. After that, the corpus of the thesis is the three empirical research chapters (chapter 5 – 7). Some overall conclusions are drawn in chapter 8.

2. Literature review

The aim of this chapter is to position the research vis-à-vis previous research efforts and to describe how the review of the literature provides the research with a well-informed perspective. Since public sector benchmarking can be seen as an approach to performance measurement and performance management, the articles selected for this literature review include the main academic contributions in the area of performance measurement, performance management and benchmarking. The review in this chapter is based on an in-depth analysis of three key journals and a width-way analysis of sub-disciplinary boundaries including another seventeen journals. It should be observed that the subsequent chapters include a wide variety of additional articles relevant to the research topic presented, not restricted by a time span.

In section 2.1 the findings of the literature review are discussed. The sections are introduced with a brief explanation of the motivation and choices made for the literature review. In section 2.2 the content of the research studies is described (2.2.1), including a discussion of the main findings (2.2.2) and the research gaps (2.2.3). There are three appendices to the literature review. The first appendix is the list of articles comprised in the review (appendix 1). The second appendix is the classification of the research questions according to Pettigrew's framework for organisational change. The third appendix is the classification of the studies according to their focus (measurement, incorporation and use) and their theoretical orientation.

2.1 A systematic review of the literature

First of all, the literature review is intended to position the research vis-à-vis previous research efforts. Section 2.1.1 explicates and motivates the choices made for the literature review. Four questions will be addressed: what is the theoretical orientation used (2.1.2); what are the methods used (2.1.3); what is the focus of the previous studies (2.1.4); and finally, what are the main research areas (2.1.5). In addition, the literature study has to provide an informed perspective. In particular, by categorizing the research questions of earlier studies, insight is gained in what the categories of variables (context, content, process) mean in PMM and benchmarking research.

2.1.1 Research method and data for the literature review

This section explicates and motivates the choices made for defining the literature review as presented in this chapter. Firstly, the review as presented in this chapter only considers journal articles. Other sources such as books and thesis are not analysed. Although the latter two sources may give a more profound insight in answering research questions, articles provide a more up-to-date picture, which allows for the inclusion of the latest findings in the review. In addition, it seems a reasonable assumption that the research issues in books and theses are comparable to those in articles, albeit with a different profundity (Van Dooren, 2006, p. 85).

Secondly, the articles were selected along two lines: in depth and width-ways. The journals for the in-depth analysis were selected based on their broad coverage of articles addressing PMM and public sector benchmarking. The three journals are: (i) *Benchmarking: an International Journal; (ii) Public Administration Review; and (iii) International Journal of Public Sector Management*. Braadbaart and Yusnandarshah (2008) assessed the evolution of public sector benchmarking. They did so with a database of 147 peer-reviewed articles published between 1990 and 2005. *Benchmarking: an International Journal* and the *International Journal of Public Sector Management* were the ones with the widest contribution of articles to the database, respectively 27 and 7 articles. The journal *Public Administration Review* has been selected since some of the key articles that aroused my interest and shaped my research ideas at the start of my research study were published in this journal (e.g., Behn, 2003, Folz, 2004, Poister and Streib, 1999). In order to surmount sub-disciplinary boundaries, the in-depth analysis was complemented by a more general search using electronic databases: EBSCO Business Source Premier and Google Scholar Advanced. The search expression included combinations of “performance”, “measurement”, “benchmarking”, “municipal”, “municipality”, “local government” and "public sector". This search resulted in articles in seventeen other journals. This extension of the literature review added articles focusing on sub-disciplinary boundaries such as organisational learning, performance information use and management accounting (see appendix 1 for the complete list of articles reviewed).

The third decision is the time span. Only articles published from 1999 to 2009 are analysed on their main contents. The period of review is overlapping with the life cycle of the benchmark Building and Housing Supervision (BWT) that was offered for the first time in 1999. The time period analysed covers one decade. The decade can be classified into two time periods. The first period (1999-2004) is overlapping with the period that Van Dooren (2006, p. 102) indicates as the period when the evaluations of New Public Management were brought out. Van Dooren observed a boom in publications in this period in comparison to the earlier time periods studied (pre-NPM era, advent of the NPM and the globalisation of the NPM rhetoric). Hypothetically, the second time period (2005-2009) can be seen as the period where the follow-up on the evaluation studies of NPM can be observed more in detail, and hence its evolution.

The fourth decision is the publication language. Only articles that are published in English have been included. Largely, this is an arbitrary decision. However, many European scholars publish in English Journals. Moreover, the performance measurement and management literature is largely English (Van Dooren, 2006, p. 88).

A total of 61 articles were selected from the in-depth analysis of the literature. The articles selected for the in-depth analysis include 22 articles from *Benchmarking: an International Journal,* 19 articles from *International Journal of Public Sector Management* and 20 articles from *Public Administration Review.* Another 39 articles were selected from the width-way analysis covering seventeen journals, i.e., *Financial Accountability and Management*, *Accounting, Organisation and Society, International Journal of Productivity and Performance Management*, *Journal of Public Administration Research and Theory* and *Public Performance and Management Review.* From the total of 100 articles reviewed, 44 articles were published in the first time period (1999 – 2004) and 56 articles were published in the second time period (2005 – 2009) identified for the literature review. The choice to include 100 articles in the literature review is a pragmatic choice.

2.1.2 What is the orientation used?

A lack of clarity exists in the literature about mixing up epistemological and ontological orientations. Easterby-Smith *et al*. (2002) contrast positivism with social constructionism, whereas Bryman (2004) contrasts positivism with interpretivism. Easterby-Smith *et al*. see the difference between positivism and social constructionism as a matter of differing ontology, whereas Bryman sees the difference between positivism and interpretivism as a matter of epistemology. Based on additional reading (Steffy and Grimes, 1986; Lee, 1991; Kim, 2003) I decided to follow the distinction as made by Bryman, which was also used in the afore-mentioned articles.

The key idea of positivism is that the social world exists externally, and that its properties should be measured through objective methods, rather than being inferred subjectively through sensation, reflection or intuition. The key idea of interpretivism is that the whole needs to be examined in order to understand a phenomenon and suggest that there are multiple realities to the understanding depending on time and place.

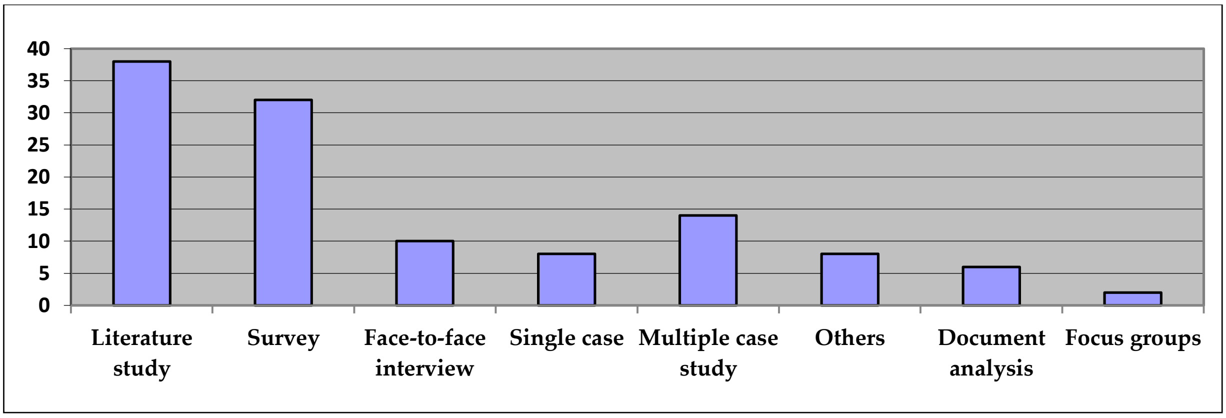
The vast majority of the studies reviewed have a positivist theoretical orientation, primarily focusing on content and process issues. Recent examples are the studies of Anand and Kodali (2008) and Ferreira and Otley (2009), respectively making an analysis of the taxonomy of benchmarking models and describing the structure and operation of performance management systems. The key methods used in the studies with a positivist theoretical orientation are literature studies and surveys. Eight out of the hundred articles have interpretative aspects (Alstete, 2008; Foster and Gallup, 2002; Kyrö, 2003 and 2004; McAdam *et al*., 2005; Parker and Bradley, 2000; Rondeaux, 2006; Wisniewski and Stewart, 2004). A recent example is the study of Rondeaux (2006) on an identity evolution taking place following the implementation of NPM principles. The author concludes that organisational identity is complex, hybrid and composite and in constant evolution according to perceptions of reality and context. Another recent example is the study by Wisniewski and Stewart (2004), who argue that performance information needs are service specific and therefore it is unlikely that there is a set of common performance measures across different services. Finally, Foster and Gallup (2002) point to the functional perspective as an important explanation of how quality and quality improvement efforts are viewed. These interpretative studies are characterised by their recognition of multiple realities and the influence of context (such as culture, organisational identity and categories of stakeholders) on PMM and benchmarking.

It can be concluded from this section, that the interpretative theoretical approach is not utilised very often in municipal benchmarking research. It is believed that municipal benchmarking research may benefit from the infusion of more interpretative elements. Studying the "human" side of a municipal benchmarking system allows for further understanding and explaining why people have different experiences with the incorporation and use of performance information.

2.1.3 What are the methods used?

Methods can explore a phenomenon by focussing on a large or a small number of observations. Large N methods are usually quantitative and rely on statistical analysis techniques such as surveys. Small N methods are usually qualitative and require more interpretation on behalf of the researcher. Some example techniques are face-to-face interviews, focus groups and case study research. Literature is seen as a non-empirical way of data collection.

Figure 2.1 Methods used in the reviewed articles



Source: the author

Figure 2.1 gives the frequencies for the methods used in the reviewed articles. Most frequently, they are literature studies (38 studies). A difference can be made between literature-based studies and pure literature studies. The conclusions of the first category deal with the theory and practice of the topics addressed and are supported by reference to the literature. For example, the article by Modell (2009b) offers a systematic review of institutional research on PMM in the public sector accounting literature. The article by Williams (2004) examines the development of performance measurement in the critical period from its origins through to 1930. The conclusions of the latter category refer to the literature itself. A total of four pure literature studies is found in the set of articles (Anand and Kodali, 2008; Braadbaart and Yusnandarshah, 2008; Dattakumar and Jagadeesh, 2003; Yasin, 2002).

The second most popular method is surveys (32 studies). Sixteen of the surveys (50%) are of local governments in the United States of America (e.g., Ammons *et al*., 2001; Berman and Wang, 2000; Cavalluzzo and Ittner, 2004; Folz, 2004; Melkers and Willoughby, 2005; Poister and Streib, 1999). Eleven surveys are of European origin, of which three are from the Netherlands. Ter Bogt (2004) published a study on the use of performance information by politicians. Van Helden and Tillema (2005) studied what the various response patterns of public sector organisations are to benchmarking. More recently, Braadbaart (2007) assessed whether collaborative benchmarking boosts the performance of public sector organisations in the Netherlands water supply industry.

Seven out of 100 studies reported to have used two different methods (Bowerman *et al*., 2002; Holloway *et al*., 1999; Longbottom, 2000; Rondeaux, 2006; Sanger, 2008; Sharifuddin bin Syed-Ikhsan and Rowland, 2004; Van Helden and Tillema, 2005).

Methods can also be classified according to their obtrusiveness. Bryman (2004, p. 545) defines unobtrusive methods as methods that do not entail the awareness among research participants that they are being studied and hence are not subject to reactivity (non-reactive). Non-participant observation and documentary sources are examples of unobtrusive methods. Whenever people know that they are participating in a study, a component of their replies or behaviour is likely to be influenced by this knowledge. This is defined as obtrusive (reactive). This is invariably the case with methods of data collection such as structured interviewing, self-administered questionnaire and structured observation.

Table 2.1 A classification of research methods

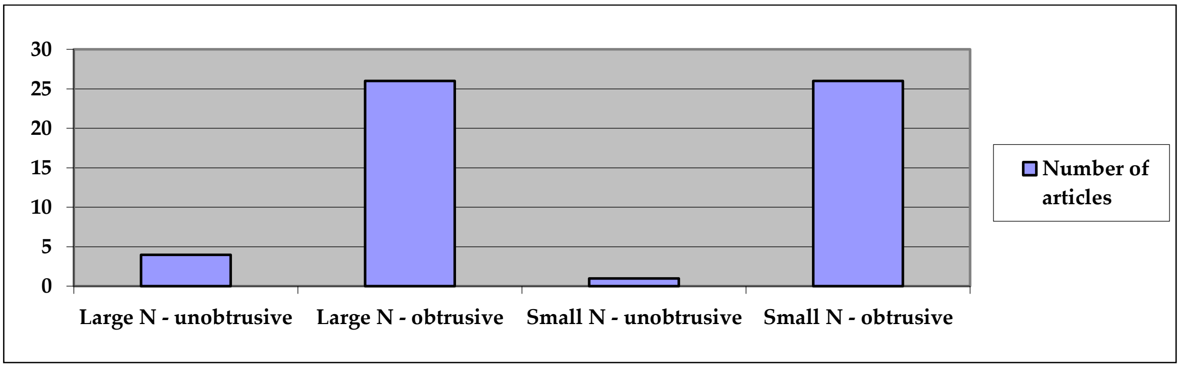
Aston University

Illustration removed for copyright restrictions

Source: Van Dooren, 2006, p. 98

Figure 2.2 below regroups the methods following the obtrusive character and the number of observations. The most popular methods are small N obtrusive (twenty-six articles) – generally case studies sometimes in combination with a small N survey (six articles). Secondly, twenty-six articles have a large N, and an obtrusive character. These are mainly the surveys. Third, three studies use unobtrusive large N. These are document analyses. The fourth method is small N unobtrusive. One article applies for this category (Rondeaux, 2006). The author uses content and speech analysis to answer the research questions (interpretative study). The small N obtrusive methods together with the literature studies add up to 39 articles.

Figure 2.2 Obtrusiveness and number of observations in the reviewed articles



Source: the author

It can be concluded from this section that although the whole array of methods is used, small N methods and literature studies are predominant. The insight that municipal benchmarking in local government agencies is a complex multi-faceted operation may explain the focus on more qualitative research approaches. The second most popular method is surveys (large N research). Surveys are used in both time periods studied (18 times in the first time slot and 14 times in the second time slot).

2.1.4 Is the focus on measurement, incorporation or use?

This analysis in this section is structured by using the classification of Van Dooren as introduced in section 1.1.4 of the previous chapter, making a clear distinction between measurement, incorporation and use of performance information. Performance measurement refers to the systematically collection of data whereas performance management refers to the incorporation and use of performance information.

Table 2.2 represents the focus of the reviewed articles. Interestingly, the combined focus on both incorporation and use has increased substantially in the second time period (2005 to 2009). A total of 75% of the studies in this time period use a combined focus on both incorporation and use against 25% of the studies in the time period between 1999 and 2004. A total of 61 studies (61%) adopt a combined focus on both incorporation and use or use in isolation in the time period between 1999 and 2009. From all hundred articles studied nineteen studies (19%) focus on incorporation, twenty-nine studies (29%) focus on use only and thirty-two studies (32%) focus on a combination of incorporation and use.

Table 2.2 Focus of reviewed articles

Focus of reviewed articles	Time period 1999 – 2004		Time period 2005- 2009		Total	
	number	%	number	%	number	%
Measurement	9	64%	5	36%	14	14%
Incorporation	12	63%	7	37%	19	19%
Use	14	48%	15	52%	29	29%
Incorporation and use	8	25%	24	75%	32	32%
Measurement and incorporation	0	0%	1	100%	1	1%
Measurement, incorporation and use	1	20%	4	80%	5	5%
Total	44	44%	56	56%	100	100%

Source: the author

The increased focus on the intertwined relation between incorporation and use in the time period between 2005 and 2009 points to the observation as made by Norman (2002) that issues of meaning are seen to be more important than measurement for the further development of the system. This observation is supported by the decline of the number of articles addressing measurement only in the period between 2005 and 2009.

2.1.5 What are the main research areas?

The analysis in this section is structured by using Pettigrew's framework for organisational change. The introduction of Pettigrew's framework for organisational change in section 1.3.1 of the previous chapter, explained that the 'how' of change can be understood from an analysis of process, while much of the 'why' of change is derived from an analysis of inner and outer context and that the 'what' of change is encapsulated under the label content. The perspective of Pettigrew on content has been taken as the basis for analysis for this chapter, and not the narrower perspective as presented in section 1.3.1 (i.e., the influence of different definitions of performance used by various institutional actors, allowing for a focus on the quality of achievement and/or the quality of actions).

Based on the review it was possible to identify twelve clusters of research areas within the three research variables on which authors have focused and where relevant research gaps were identified.

An overview of the twelve clusters is given in table 2.3. The research areas are linked to the three research variables: process, context and content. The numbers relate to the number of studies found per cluster and research variable. It can be observed that some articles fit in more than one cluster. The articles have been plotted to the clusters based on the key emphasis of the article.

Table 2.3 Clusters of research areas

Total = 100 studies	Process (how) (35)	Context (why and when) (33)	Content (what) (32)
Clusters of research areas	1. Dysfunctional effects and consequences (7)	1. Organisation's environment (outer context) (9)	1. Quality of PMS / municipal benchmark (15)
	2. Data purpose and use (16)	2. Organisational factors (inner context) (13)	2. Benchmarking partners (4)
	3. Compatibility of uses of performance information (5)	3. Theoretical approaches (9)	3. Involvement of employees and managers (4)
	4. Implementing a change process (7)	4. Transferability of private sector approaches (2)	4. Managerial effectiveness (9)

Source: the author

2.2 Content of the systematic review

The twelve clusters of research areas identified and presented in table 2.3 are used as a basis for discussing the content of the research studies in the next section. In section 2.2 the content of the research studies is described (2.2.1), including a discussion of the main findings (2.2.2) and the research gaps (2.2.3).

2.2.1 Description of the content

The <u>first</u> cluster of research areas relates *process* to the impact of PMM and municipal benchmarking on organisational learning and performance. Some issues and exemplary studies are discussed below.

First, a repeated issue concerns the **dysfunctional effects and consequences** of performance measurement. It is argued by Moynihan (2006) that the managing for results doctrine has been only partially adopted. Governments selected some of the NPM ideas but largely ignored others (i.e., enhancing managerial authority). In addition to this Newcomer (2007) observes a lack of clarity in expectation among public managers regarding how performance measures may be used. The success of performance measurement systems in the public sector can be explained (De Bruijn and Van Helden, 2006) if these systems are developed and used in an interactive way between managers and professionals. Without a behavioural-oriented approach, systems of performance management can create strong incentives for perverse behaviour and might therefore be a victim of the Law of Decreasing Effectiveness: more control, appraisal and sanctions will lead to more perverse effects and reduced effectiveness. Several dysfunctional effects of performance measurement can be identified in the literature (De Bruijn, 2002; Pidd, 2005; Van Thiel and Leeuw, 2002); such as, performance measurement adds to internal bureaucracy, blocks innovations and ambitions, prompts game playing and kills system responsibility. To counteract these effects, performance assessment systems requires among others the use of multiple indicators, referring to different aspects of policy implementation (tangible and non-tangible), reflecting the interests of all stakeholders and using a process perspective of performance measurement in addition to a product perspective (De Bruijn, 2002; Van Thiel and Leeuw, 2002). The authors referred to in this cluster indicate several dysfunctional effects and consequences of performance measurement. A list of scattered approaches to counteract the dysfunctional effects could be distilled, among others the use of a behavioural-oriented approach.

The most studied variable of the articles reviewed is **data purpose and use**. Managers have an internal perspective, whereas politicians seem to have a citizen perspective and a financial perspective on performance. The citizen perspective is typical of governmental organisations and it can limit the implementation of more business-like ways of running such organisations (Jansen, 2008). On the whole politicians do not take much interest in performance information, and neither do citizens, unless and until disasters, scandals or breakdowns come along (Pollitt, 2005). Based on a study among 698 aldermen of Dutch municipalities with 20,000 or more inhabitants (Ter Bogt, 2004), it can be indicated that in general Dutch aldermen see little value in the output-oriented performance information that is available in the planning and control documents of their organisations and they use it only infrequently. Aldermen seem to prefer qualitative, rich information to historical, standard, written, and numeric output information in planning and control documents.

Several authors mention possible explanations for not using performance information and solutions to counteract these. Explanations for not using performance information mentioned in the articles studied (Pollitt, 2006; Wisniewski and Stewart, 2004; Greiling, 2005; Pollanen, 2005) include: timing of evaluations and performance reports, high volumes of information, timeliness, credibility, limited thought as to the most appropriate way of reporting performance, little has been done to assess stakeholder satisfaction with either the performance information provided or the way it has been provided and ambiguity of performance measures. Adcroft and Willis (2005) refer to six systemic problems of using performance measurement for performance improvement. I would like to highlight two systemic problems to which the authors refer:

(i) The more services are broken down and deconstructed into ever smaller components, the less the performance of the whole service is being measured;

(ii) The presentation of performance measurement in league tables assumes that all those being measured start from the same point. There are usually more sources of difference between same-service providers than there are similarities.

Some solutions mentioned in the literature to improve the use of performance information are (Tillema, 2007; Ammons and Rivenbark, 2008; Julnes and Holzer, 2001): having a benchmarking culture and powerful managers, cooperation between public sector organisations in the performance improvement stage, the incorporation of performance measures into key management systems, collection of and reliance on higher-order measures (efficiency measures, rather than simply workload or output measures), conduct an assessment of organisation's "readiness" to develop and implement performance measures and identify and involve the organisation's internal and external interest groups. In addition to this Moynihan and Ingraham (2004) observe that leaders who become publicly involved in, and identified with, setting strategic goals and demanding performance information have a positive effect on encouraging employees to take performance measurement seriously as a process, and to incorporate performance information into decision making. The evidence suggests, therefore, that the influence of leadership will vary with the type of leader promoting the system, and the level of employee responding to leader initiatives. In addition to this it is observed (Nicholson-Crotty *et al*., 2006) that the choices managers make about measurement can have a significant impact on their evaluations, including assessments of whether their organisation has a problem.

It can be concluded from this research area that the different perspectives on performance in the public sector affect the extent to, and the way in which, performance information is used. In addition to this several problems for not using performance information and solutions to counteract these can be found in literature.

Next, some studies explore the **compatibility of uses of performance information**. The benchmarking method should facilitate organisations' responses to calls from a diversity of interested parties. Wynn-Williams (2005) studied how benchmarking can help to provide meaningful and relevant information to funders, service providers, service recipients and other interested parties. The study proposes that a combination of internal benchmarking, process benchmarking and increased public documentation will enhance reporting systems in any public sector organisation. Van Helden and Tillema (2005) argue that some factors influencing the willingness of organisations to take part in a benchmarking project and the resulting performance improvement actions may conflict with others. The authors state that if, for instance, economic gain motives are strongly influencing organisations to take part in a benchmarking project the resulting conformist behaviour may conflict with their desire for autonomy. It can be concluded from the articles reviewed in this cluster that the theoretical and conceptual rift running through the literature (benchmarking as a vehicle for competition or cooperation) influences the compatibility of uses of performance information.

Finally, several scholars pointed to the requirement of **implementing a change process** leading to and supporting strategic decisions based on the municipal benchmark. Coplin *et al*. (2002) mention that academic and professional publications give the impression that performance measurement is a growing government practice, in actuality the use of this technology is not as deep and widespread as it may appear. The authors suggest that the most important danger of benchmarking may be neither its 'perverse' effects nor its abuse by central authorities to retain control. The real danger of benchmarking may rest with the fact that even in the situation where the benchmarking analysis is initiated, designed and conducted by the benchmarking organisations themselves, they are caught in a game of incessant organisational change that some of the participants will always lose. Benchmarking depends on the freedom or self-governing capacities of those who are benchmarked (Triantafillou, 2007). Fernandez and Rainey (2006) suggest eight factors to which change managers and change participants pay attention in order to successfully implement a change process: (1) ensure the need; (2) provide a plan; (3) build internal support for change and overcome resistance; (4) ensure top management support and commitment; (5) build external support; (6) provide resources; (7) institutionalise change and (8) pursue comprehensive change.

Jones (1999) describes the case study of Wollongong City Council (Australia), for the introduction and use of benchmarking as part of a quality-oriented cultural change program. From the case study it can be concluded that comparative benchmarking provides the trigger (by identifying areas of poor performance), but process knowledge provides the means of identifying how performance can be improved. The authors referred to in this cluster suggest that professionals need to function as "change agents", using a variety of strategies to gain acceptance and understanding of the strengths and limitations of performance measurement.

The second cluster of research areas relates *context* to the impact of PMM and municipal benchmarking on organisational learning and performance. Some issues and exemplary studies are described below.

The **organisation's environment** (outer context) may influence PMM and municipal benchmarking. Holloway *et al*. (1999) argue that benchmarking is only as effective as the people who apply it. In addition to this the author observes that compatibility with the outer context in which it is used is important. PMM instruments such as municipal benchmarking are not merely about instrumental change, but about changes in identities that surround public organisations. Organisational identity is complex, hybrid and composite and in constant evolution according to perceptions of reality and context (Rondeaux, 2006; Van Bockel and Noordegraaf, 2006). Talbot (2008) indicates that it is the totality of a performance regime which potentially shapes or steers performance for specific organisations rather than the narrow principal-agent assumptions often made about performance drivers. This last observation is in contradiction with the viewpoint of Andersen *et al*. (2008), who demonstrate that the principal-agent theory suitably describes the context within which compulsory benchmarking can be put to useful use. The articles reviewed in this cluster confirm that the outer contextual factors play an important part both in establishing a need to use approaches such as benchmarking, and in encouraging commitment to their use.

Organisational factors (inner context) can be found in Amaral and Sousa (2009). They made a distinction between different barriers to benchmarking initiatives: organisational barriers (people, culture and context), benchmarking project management barriers (planning and implementation and leadership) and benchmarking data barriers (difficulty to access / compare data). Berman and Wang (2000) find that the success of performance measurement is greatly affected by underlying organisational capacities. Their findings are based on a national survey among U.S. counties with populations over 50,000.

Capacity requires that organisations are able (1) to relate outputs to operations; (2) to collect timely data; have (3) staff capable of analysing performance data; (4) adequate information technology; and support from (5) department heads and (6) elected officials. In addition to this Cavalluzzo and Ittner (2004) find that organisational factors, such as top management commitment to the use of performance information, decision-making authority, and training in performance measurement techniques, have a significant positive influence on measurement system development and use. Foster and Gallup (2002) point to the differences that exist in how people with different functional job classifications view quality and quality improvement efforts. They propose the use of cross-functional teams to improve quality. Several authors (Longbottom, 2000; Mwita, 2000; Teelken, 2008) emphasise the need to link performance measurement systems design with issues of policy, strategy, operations, assessments and information systems. The implementation of such systems seems to occur outside the primary process of the organisation. According to Van Dooren (2006) a performance measurement policy is often the missing link in public sector reform. It is as well suggested (Mwita, 2000) that management accounting and other performance measurement practices need to be evaluated not just from an economic perspective, but also from a social, behavioural and managerial perspective, within an overall organisational context. According to Teelken (2008) institutional as well as professional theories supplement each other in a fruitful way in order to explain the difficulties with implementation of PMS. The 'human' side of PMS should be able to acknowledge and bridge the gap between the primary process and the organisation of performance measurement. While institutional theory helps to understand the initial resistance to change, professional theory helps to understand the pragmatic embrace by the individual. Broadbent and Laughlin (2009) developed a model of the alternative nature of PMS lying on a continuum from 'transactional' (driven by the exercise of instrumental rationality) at one end to 'relational' (driven by the exercise of communicative rationality) at the other, built on respectively underlying instrumental and communicative rationalities and guided by a range of contextual factors. Finally, Askim *et al*. (2007) observe that factors such as network and administrative characteristics and management and political participation are found to influence organisational learning from benchmarking. The authors have come to this conclusion by studying a nationwide

Norwegian benchmarking project (2002-2004) for local governments in which more than 300 municipalities took part, grouped in 40 benchmarking networks. The articles reviewed in this cluster confirm that the inner context of an organisation has a significant influence on measurement system development and use.

Theoretical approaches to benchmarking have been identified among others by Yasin (2002), who reviewed the literature related to benchmarking practice and theory from 1986 to 2000. The study revealed that the earlier stages of benchmarking literature stressed a process and/or activity orientation. In later stages, the scope of benchmarking literature appears to have expanded to include strategies and systems. Despite these advancements, the field of benchmarking still suffers from the lack of theoretical developments. In 2009, Moriarty and Smallman support this view by observing that benchmarking remains theoretically underdetermined, with publications focusing on pragmatism and praxis rather that epistemology. They suggest establishing a theoretical basis for benchmarking based on classical and modern theories of causation in conjunction with economic welfare theory to quantify the advancement from a current to superior state of affairs. Insights of relevant organisational learning literature on results-based reforms (Moynihan, 2005) shows that most results-based reforms target narrow process improvement (single-loop learning) rather than a broad understanding of policy choices and effectiveness (double-loop learning). Brignall and Modell (2000) argue that studies of managerial choice constitute a useful starting point for analysing how PM practices change in highly institutionalised settings, such as the public sector. Ter Bogt (2008b) indicates that an institutionalist perspective enables one to study change processes in organisations and to observe issues and developments that might not be noticed when a more functional and short-term perspective is chosen. His research findings suggest that social factors and structures influence the accounting change process in the organisations to a considerable extent. NPM is impacting on public officials as their roles and the work they do, the ways in which they are managed, their relationships with the public and the criteria by which they are assessed, both internally and externally, are continually evolving. It is interesting to note that Horton (2006) observes that, in spite of all the changes, much of the traditional public administration cultures remain. The article also highlights the contribution that cultural and social theories, drawn from anthropology and organisational psychology, make to an understanding of the processes by which public servant's identity are formed and changed. Johnsen and Vakkuri (2006) suggest that in the Nordic countries (and maybe also the Netherlands) performance measurement may be used relatively more for dialogue and learning than for management control. Because the Nordic model depends on economic flexibility, social innovation, and political compromises the Nordic perspective may favour a homeostatic model more than a cybernetic model. It can be concluded from the articles studied in this cluster that the field of benchmarking still suffers from a lack of theoretical developments, remaining theoretically underdetermined, with publications focusing on pragmatism and praxis.

A different subject related to context as an explanatory variable is the **transferability of private sector approaches** to the public sector. Benchmarking is one of the imports from the private sector. Whereas, the public sector has wholeheartedly embraced benchmarking, it is not obvious how benchmarking affects public sector organisations. Braadbaart (2007) refers to two disparate scientific literatures. One school of thought presents benchmarking as a tool that helps national governments control public service providers (compulsory benchmarking based on competition). A second literature envisages benchmarking as a vehicle for collaborative learning among public sector organisations (voluntary benchmarking based on cooperation). Whereas the voluntary benchmarking literature is sceptical of the prospects for compulsory benchmarking, arguing that a top-down imposition of benchmarking may inhibit learning and aggravate rather than solve problems of performance assessment, the compulsory benchmarking literature asserts that performance will only improve under compulsory benchmarking imposed by a regulator. The study of Braadbaart addresses the question whether voluntary benchmarking can boost the performance of public sector organisations. He presents evidence from the Dutch water supply industry focusing on transparency and performance in voluntary benchmarking by applying a quasi-experimental method to 1989-2000 time series data on benchmarking and non-benchmarking water utilities. It was concluded in the study that benchmarking immediately enhanced transparency, but only affected utility economic performance after benchmarking information entered the public domain. A decade before Kouzmin *et al*. (1999) made a preliminary attempt at drawing some conclusions about expanding German and other European benchmarking experiences. According to the authors benchmarking can be seen as a learning strategy. The learning-effects of benchmarking are, to a very high degree, dependent on adequate organisational conditions and managerial solutions. In their view, a full utilisation of the "learning" potential of benchmarking is possible only if elements of competition and cooperation are combined and cogently managed. The authors refer to the paradox that the outstanding pre-condition for learning in organisations is the creation of the "learning" organisation in structural and cultural terms. The articles studied in this cluster indicate that authors are divided which element(s) will improve performance: competition (compulsory), cooperation (voluntary) or a combination of both.

The third cluster of research areas relates *content* to the impact of PMM and municipal benchmarking on organisational learning and performance. Some issues and exemplary studies are noted below.

The **quality of PMS**, and more specifically a municipal benchmark, is a noteworthy research concern. Several authors argue that the benchmarking technique has seen a steady growth and evolved into a mature and strongly international field of research (Braadbaart and Yusnandarshah, 2008; Dattakumar and Jagadeesh, 2003). Nevertheless Fryer *et al*. (2009) state that the expected improvements in performance, accountability, transparency, quality of service and value for money have not yet materialised in the public sector. The authors identified three classes of problems with performance management in the public sector: (i) technical problems (relate to the indicators and the data, their collection, interpretation and analysis); (ii) systems problems (integrating performance systems with the existing systems, a lack of strategic focus which encourages short-termism, the ambiguity of performance objectives, sub-optimisation and the cost of performance management); and (iii) involvement of employees and managers. Ferreira and Otley (2009) put forward a performance management systems framework as a research tool for describing the structure and operation of performance management systems in a more holistic manner, serving the quality of a PMS.

Behn (2003) argues that without at least a tentative theory about how performance measures can be employed to foster improvement, public managers will be unable to decide what should be measured. As part of their overall management strategy, public managers can use performance measures to evaluate, control, budget, motivate, promote, celebrate, learn and improve. The management purposes can be classified along the theoretical and conceptual rift running through the literature (benchmarking as a vehicle for competition or cooperation) as referred to earlier. Hinton *et al*. (2000) note that a great deal of benchmarking activity can be described as "results" benchmarking as opposed to "process" benchmarking. The authors strongly recommend that benchmarking includes an examination of the underlying processes. In addition to this it is suggested that bench markers pay at least as much attention to the organisational climate as to the technical or formal steps taken. This view is supported by Williams (2004) who argues that history shows that performance measurement does not refer to a particular empirical technique. Instead, it refers to the application of relevant techniques to the problem of observing government at work (delivery of public services). Norman (2002) makes a distinction between three different users of performance measurement systems. The *true believers* think that more effort should be put into creating clearer, more observable measures that emphasise outcomes. *Pragmatic sceptics* see reported measures as part of a new game of public management and at best a starting point for asking about the substance behind the form. *Active doubters* believe that too much emphasis on measurement gets in the way of the 'real work' of developing relationship-based work in a political environment. Norman suggests that issues of meaning are seen to be more important than measurement for the further development of the system. The articles referred to in this cluster suggest that a PMS framework should include an examination of the underlying processes and the inner context of the organisation to support its quality.

Next, some studies explore the role that **benchmarking partners** play in creating value by being involved in performance comparison projects. According to Ammons *et al*. (2001) participating units that expect obvious and easy solutions for inefficiencies and service shortcomings face inevitable frustration. Successful bench markers make plans for the *use* of the tool, and they carry out those plans. Their less successful counterparts often incorrectly assume that good results will accrue simply from *having* the tool. Dawes *et al*.

(2009) observe that network development processes that emphasise early, open dialogue and examination of assumptions and expectations do better than those that rush forward with a fixed IT solution in mind. The studies suggest that network builders, political leaders and public managers need to invest in developing public management skills to add value in building public sector knowledge networks.

The **involvement of employees and managers** in benchmarking practices is described. Alstete (2008) states, that there is a misunderstanding among employees regarding performance measurement and performance management. The majority of the participants in his research study believed that benchmarking was merely obtaining the comparable data (performance measurement). They did not realise that in order to truly benchmark, there should be a plan that is implemented and continuously studied for on-going improvement. The author reveals that a more precise terminological use of benchmarking practices should be promoted and used to increase involvement. It is found that most benchmarking efforts are hampered by resistance from both managers and lower level employees to change (Goncharuk and Monat, 2009; Streib and Poister, 1999). Streib and Poister (1999) learned that lower level employees are typically not involved in the development of performance measures. Goncharuk and Monat (2009) suggest that benchmarking efforts could be enhanced by integrating employee motivation/behaviour programs with the benchmarking efforts. A multiple case study including six cities in the United States implemented by Sanger (2008) revealed that the jurisdictions and agencies with the best performance reporting and performance management efforts have strong mission-driven leaders at the helm who communicate the mission, motivate employees, shape strategies, and provide support, rewards, and sanctions for achievement. The articles studied in this cluster suggest that involvement of employees and managers is positively associated with improved performance.

A fourth, topical issue in the relation between content and impact of PMM and municipal benchmarking on organisational performance is **managerial effectiveness**. According to Willcocks (2002) managerial effectiveness is essentially about understanding, reinterpreting and making sense of different role expectations. Not just meeting the expectations of others (single-loop learning), but proactively challenging and influencing a range of different expectations (double-loop learning). Willcocks (2002) offers a multi-theoretical and multi-level framework that focuses upon different levels of public sector effectiveness (including managerial effectiveness). Modell (2009b) observes that research is beginning to move beyond simplistic portrayals of PMM as decoupled or loosely coupled practices to pay greater attention to how it is implicated in the formative stages of institutional processes and the role of agency exercised at different levels of analysis. Vigoda-Gadot and Yuval (2003), show that managerial quality leads to administrative performance and ultimately to trust in governance. The findings support assumptions that administrative performance may be treated as a precondition to trust in governance rather than trust serving as the precondition to performance. It can be concluded from the articles in this cluster that different levels of public sector effectiveness (individual, managerial, organisational and inter-organisational) are recommended by the authors for studying managerial effectiveness.

2.2.2 Main findings

The review of the studies on performance measurement and management in general and benchmarking in particular demonstrates the many aspects of research on the subject.

The first cluster of research areas related *process* to the impact of PMM and municipal benchmarking on organisational learning and performance. The main findings that emerged from the review of the articles in this cluster are:

- Authors referred to several dysfunctional effects and consequences of performance measurement. A list of scattered approaches to counteract the dysfunctional effects could be distilled, among others the use of a behavioural-oriented approach.
- The different perspectives on performance in the public sector affect the extent to, and the way in which, performance information is used. In addition to this several problems for not using performance information and solutions to counteract these can be found in literature.

- The theoretical and conceptual rift running through the literature (benchmarking as a vehicle for competition or cooperation) influences the compatibility of uses of performance information.
- Professionals need to function as "change agents", using a variety of strategies to gain acceptance and understanding of the strengths and limitations of performance measurement.

The second cluster of research areas related *context* to the impact of PMM and municipal benchmarking on organisational learning and performance. The main findings that emerged from the review of the articles in this cluster are:

- Outer contextual factors play an important part both in establishing a need to use approaches such as benchmarking, and in encouraging commitment to their use.
- The inner context of an organisation has a significant influence on measurement system development and use.
- The field of benchmarking still suffers from a lack of theoretical developments, remaining theoretically underdetermined, with publications focusing on pragmatism and praxis.
- Authors are divided which element(s) will improve performance: competition (compulsory), cooperation (voluntary) or a combination of both.

The third cluster of research areas related *content* to the impact of PMM and municipal benchmarking on organisational learning and performance. The main findings that emerged from the review of the articles in this cluster are:

- The content of a PMS framework should include an examination of the underlying processes and the inner context of the organisation to support its quality.
- The studies suggest that network builders, political leaders and public managers need to invest in developing public management skills to add value in building public sector knowledge networks.

- Involvement of employees and managers is positively associated with improved performance.
- Different levels of public sector effectiveness (individual, managerial, organisational and inter-organisational) are recommended by the authors for studying managerial effectiveness.

2.2.3 Research gaps

Three key research gaps and possibilities for further research into the impact of PMM and benchmarking on organisational learning and performance in the public sector emerged from the literature review.

The first cluster of research areas related *process* to the impact of PMM and municipal benchmarking on organisational learning and performance. As referred to by Pettigrew (1987, pp. 657-658), *"The process of change refers to the actions, reactions, and interactions from the various interested parties as they seek to move the organisation from its present to its future state"*. The main research gap that emerged from the review of the articles in this cluster is that studies on the factors that affect the operation of performance measurement within a government agency have the character of laundry lists (e.g., the listing of dysfunctional effects and reasons for not using performance information). They do not search for underlying classifications and mechanisms related to the different uses of performance information.

The second cluster of research areas related *context* to the impact of PMM and municipal benchmarking on organisational learning and performance. Pettigrew observes (1987, p. 657) that "*outer context refers to the social, economic, political and competitive environment in which the organisation operates. Inner context refers to the structure, culture, and political context within the organisation through which ideas for change have to proceed*." Although it can be observed from the review that the inner context of an organisation has a significant influence on measurement system development and use, these studies do not address the influence of the political context on benchlearning.

The third cluster of research areas related *content* to the impact of PMM and municipal benchmarking on organisational learning and performance. Content is concerned with the areas of transformation and the tools and techniques used to effect change. The content of benchlearning in this thesis refers to the influence of different definitions of performance used by various institutional actors, allowing for a focus on the quality of achievement and/or the quality of actions. This specific aspect of content has not been addressed in the articles reviewed as part of this cluster and can be identified as the third main research gap.

The next chapter presents the philosophical and methodological base of the research including some critical choices regarding the research strategy, design and methods that best suit the themes under investigation.

3. Philosophical Perspective and Methodology

This chapter presents the philosophical and methodological bases of the research. It starts by presenting two different research strategies: quantitative and qualitative research. In the first section critical choices are provided regarding the research strategy, design and methods that best suit the themes under investigation. The research questions of this research study are presented in section 3.2. By consideration of the data required a multiple-case study approach is selected for this research study, because of the richness and depth of the evidence required (see section 3.3). A description of how the case organisations were selected is provided in section 3.4. In section 3.5 the data collection methods that were used are described in more detail. In the subsequent section (3.6) a description is given of the data analysis methods used. In section 3.7, an overview is given of the profile of the interviewees. The full list of interviewees can be found in appendix 4.

3.1 Critical choices

Whereas quantitative research has its roots in the positivist paradigm, it is often assumed in management accounting research that qualitative studies are based on subjectivism. Ahrens (2008) indicates that "interpretative" is often used interchangeably with "qualitative" to characterise a study's methodology, or general approach to studying. Quantitative research entails a deductive approach to the relationship between theory and research. This approach represents the view that hypotheses (drawn from existing theory) must be subjected to empirical scrutiny. Qualitative research tends to be inductive in nature, with the emphasis on the generation of theories (Bryman, 2004). One important feature of quantitative techniques is that the process of data collection is distinct from data analysis, whereas with qualitative techniques this is a continuous, iterative process (Easterby-Smith et al., 2002). Methods that are typically associated with quantitative research include: structured interviews, surveys, tests and structured observation. Methods that are typically associated with qualitative research include: focus groups, observation and non-structured interviews.

Another view on reasoning and developing explanations is the process of abduction. Abduction is about developing theoretically informed explanations to new empirical observations relying on the skilful development of theoretical explanations with the help of everything that is known empirically and theoretically about the issue being examined (Lukka and Modell, 2010; Modell 2009a). Modell (2009a, p. 213) observes the following:

> *"Abduction differs from the hypothetic-deductive mode of analysis, prevalent in functionalist management accounting research, as well as the strongly inductive approach found in much interpretive research. In contrast to the hypothetic-deductive mode, pivoting on empirical testing of hypotheses derived from extant theories, abduction is about developing theoretical explanations based on emerging empirical observations. However, it does not move directly from empirical observations to theoretical inferences, as is the case in purely inductive research, but relies heavily on theories as mediators for deriving explanations".*

Recent studies (Lukka and Modell, 2010; Modell, 2009a; Kakkuri-Knuuttila *et al*., 2008; Ahrens, 2008) indicate that interpretative research in management accounting combines subjectivist and objectivist features. Kakkuri-Knuuttila *et al*. (2008) are very explicit in this by indicating that strict distinctions between objective and subjective approaches to research make no sense. The authors claim that interpretative studies often present causal and explanatory hypotheses. In their study the authors explicate how concepts from different paradigms, such as interpretations, understanding meanings, and causality, can successfully co-exist and co-operate within a single study. The archetypical subjectivist position is based on the ontological assumption that reality is largely a projection of human mind. The main emphasis is on the meaning people attach to phenomena. Lukka and Modell (2010) indicate that interpretative research tends to entail highly context- and time specific analyses of how people communicate and act in a particular social setting. The interpretative approach tends to favour the 'emic' perspective (native insider) rather than the 'etic' perspective (outsider) (Kakkuri-Knuuttila *et al*., 2008; Lukka and Modell, 2010). In their paper Kakkuri-Knuuttila *et al*. (2008) observe that an interpretative study without the mobilisation of the objectivist dimension is bound to be viewed as a relatively uninteresting descriptive summation of interpretations developed by the examined actors.

Ahrens and Chapman (2006) discuss the ways in which the doing of qualitative research brings to bear discipline on the researcher allowing outsiders to assess the trustworthiness of their accounts. The authors argue that theory, method, methodology, and knowledge gains in qualitative field studies must achieve 'fit' in order to contribute to the literature. Since the paradigm is the foundation of the research design and the source of new potential outcomes, it becomes a strategic decision. Easterby-Smith *et al*. (2002, p. 43) suggest some key choices to take into account before selecting a paradigm. Table 3.1 summarises the most critical choices in selecting a research posture.

Table 3.1 Paradigm's influence on key choices of research design

	Positivism	**Interpretivism**
Researcher's role	Researcher is independent	Researcher is involved
Sample	Large samples	Small numbers
Theories	Testing theories	Generating theories
Methods used	Experimental design	Fieldwork methods
Level of generality	Universal theory	Local knowledge

Source: Adapted and modified from Easterby-Smith *et al*., 2002; Kennerley *et al*., 2004.

The literature study as presented in the previous chapter showed that small N methods, literature studies and surveys are predominant in PMM and benchmarking research in the period between 1999 and 2009. Much of the academic research reported takes a positivistic approach. The inherent problem with this approach is that the results provide little insight to managers and practitioners which might help them identify and understand the factors and system requirements that strengthen the relationship between municipal benchmarking and performance. Previous research suggests that municipal benchmarking research may benefit from the infusion of more interpretative elements.

The descriptive, explorative and more qualitative oriented nature of case study methodology is chosen to describe and interpret the information required to answer the research questions of this research study. The cases are basically intended to describe and explore the existing situation and to refine existing theory. Existing theory is addressed as a starting point, and is allowed to unfold gradually through interaction with data. By adopting this approach, theoretical explanations are developed, based on emerging empirical observations. As with abduction, this research study relies heavily on theories as mediators for deriving explanations.

3.2 Research questions

As emerged from the literature review, studies that relate *process* to the impact of PMM and municipal benchmarking on organisational learning and performance have the character of laundry lists. They do not search for underlying classifications and mechanisms related to the different uses of performance information. Although it is observed that the inner context of an organisation has a significant influence on measurement system development and use, studies relating *context* to the impact of PMM and municipal benchmarking on organisational learning and performance do not address the influence of the inner political context on benchlearning. The *content* of benchlearning in this thesis refers to the influence of different definitions of performance used by various institutional actors. This specific aspect of content has not been addressed in the articles reviewed and could be identified as the third main research gap.

Therefore, this research aims to provide deeper insights by studying these topics in the context of municipal organisations in the Netherlands. The following six empirical questions are addressed by this research study:

1. Which levels of analysis, learning processes, learning barriers and supporting factors can be observed in municipal benchmarking in the building and housing sector in the Netherlands?
    a. In what way can the 4i framework of organisational learning be adjusted to provide a helpful starting point for studying municipal benchmarking?

b. What empirical observations can be made when applying the adjusted learning framework to the four case studies?

2. Which forms of strategic stances can be seen in the four cases?

3. Which forms of power can be seen in the four cases?

4. How do different forms of strategic stances and power affect the learning outcome of benchmarking in the building and housing sector in the Netherlands?

5. Who are the main groups of stakeholders in the building and housing sector influencing municipalities and how do they understand performance?

6. What is the influence of different definitions of performance by various institutional actors on benchlearning?

The first research question is addressed in chapter 5. Questions two, three and four are addressed in chapter 6 and the last two research questions are addressed in chapter 7.

3.3 Research methods

The research questions formulated require a deep insight into the process, context and content of benchlearning. To answer the research questions, it is necessary to identify and understand the factors and requirements that strengthen the relationship between municipal benchmarking and performance.

A multiple-case study approach is selected for this research study, given the richness and depth of the evidence required. A case study is described by Yin (1994, p. 13) as "an empirical enquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between the phenomenon and its context are not clearly evident".

In order to address the identified issues of the research, four case studies in the building and housing sectors of municipalities were carried out. The case study is a research strategy that focuses on understanding the dynamics of management and organisational processes. A case study entails the detailed exploration of a specific case through a wide variety of data collection techniques. The following data collection techniques were used: unstructured interviews (pilot interviews), document analysis and semi-structured interviews. The collection of multiple data sources facilitates synergy and the identification of interrelationships (Eisenhardt, 1989, p. 538).

3.4 Selection of case studies

The cases selected for this research study are purposive. Babbie (2010) explains that the purposive case selection includes those cases that the researcher believes will yield the most comprehensive understanding of the subject of the study, based on the intuitive "feel" for the subject that comes from extended observation and reflection. In this research, the purposive case selection was not based on intuitive feel. Instead, it was assumed that the best way to get answers on the research questions is to include those cases that have a certain level of maturity with benchmarking in general and the Building and Housing Supervision benchmark in specific. In the course of time it has been found that the benchmark results between small and large municipalities differ in terms of throughput, responsiveness and efficiency (Hoogwout, 2007). For this reason, the size (number of inhabitants) of the particular case has been added as selection criteria.

Pettigrew (1997, p. 9) mentions the importance of judging when a process begins and ends. This study is designed around two time periods: (1) an appropriate time interval in which to collect data in order to see effects. Bourne *et al*. (2000) observe that the process of fully implementing a PMS in an organisation takes 13 to 18 months; and (2) an appropriate time period in which major change effects would be seen. Hinings (1997, p. 499) observes that existing literature indicates that major change takes between three and ten years. An overview of the selection criteria is given below:

- The size (number of inhabitants) of the particular case (municipalities were selected between 100,000 and 250,000 inhabitants to support comparability);
- The maturity of the selected benchmark within the particular case (municipalities were selected that participated at least since 2005);
- The amount of data and analysis already available (municipalities were selected that participated three or more times in the benchmark cycle between 1999 and 2009);

A total of 43 municipalities participated in the benchmark in the period between 1999 and 2009. Of these, a total of 21 municipalities have less than 100,000 inhabitants and 4 municipalities have more than 250,000 inhabitants. 18 municipalities meet the selection criteria having between 100,000 and 250,000 inhabitants. Of these 18 municipalities a total of 12 municipalities participated since 2005. Of these 12 municipalities 5 municipalities participated three or more times in the benchmark cycle between 1999 and 2009 (Almere, Breda, Eindhoven, Emmen and Groningen).

The municipality of Almere is one of the fastest growing municipalities of the Netherlands. The initial impetus for construction was given in 1975. The idea is that the municipality of Almere expands to a city of 350,000 inhabitants in 2030, possibly becoming the fifth city in the Netherlands in terms of population size. It is expected that the major investments in the construction and housing sector will influence their benchmark results. Because of this consideration, the following four case municipalities have been selected for this research study: Breda, Eindhoven, Emmen and Groningen. It is believed that these cases yield the most comprehensive understanding of the subject of the study.

3.5 Collection of data

In this study, twelve unstructured interviews were the first step, conducted between December 2008 and September 2009. Unstructured interviews tend not to use prepared questionnaires or interview schedules; they rather have a number of themes or issues that they aim to explore. These unstructured interviews included introductory meetings with three of the four case municipalities and other relevant organisations. The purpose of the first round of data gathering was threefold. The first objective was to gain insight into the several benchmarks offered to municipalities and to select the benchmark that would be the focus of this research study. The second objective was to understand the boundaries between the phenomenon and the context in order to establish the topics for the semi-structured interviews in the second round of data gathering. The third objective was to gain insight in the terminology used within the building and housing sector. This was a necessary condition for developing the interview guideline and preparing for the semi-structured interviews.

Semi-structured interviews were the second step in this research. A total of 24 semi-structured interviews were conducted between November 2009 and September 2010. The purpose of the second round of data gathering was threefold. The first objective was to gain a better understanding of the underlying classifications and mechanisms related to the different uses of performance information. The second objective was to explain context-specific tendencies. The third objective was to understand the constructs that the interviewee uses as a basis for opinions and beliefs about the impact of benchmarking on organisational learning and performance. In face-to-face meetings interviewees were given the opportunity to reflect on the impact of municipal benchmarking on organisational learning and performance within their municipality. The interview guide used during the semi-structured interviews was structured around the three areas of particular focus for this research study (process, context and content). Semi-structured interviews allow for easier coding and analysis.

The interviews were complemented by a document analysis to investigate the performance changes between the different years in the four case municipalities. The documents used for this purpose included the reports that present and discuss the benchmarking results covering the period 1999 – 2009. Additional documents (such as strategy reports, minutes, activity plans, etc.) collected during several meetings were also used. Reference to these documents is made in the empirical chapters (chapter 5 to 7).

Prior to the semi-structured interviews, interviewees were invited to be involved in the research project with an e-mail briefly explaining the purpose, nature, and the prospective contribution of the research outcome. It was clearly explained that the research study was part of a PhD project. The informed consent of the participants has been obtained prior to each interview indicating the confidentiality of the information. In addition to this expectations and procedures of the interview session were explained. In order to obtain authentic attitude reactions it was considered important not to provide any interview questions prior to the interview. Interviews were held in the office of the respondent in the Dutch language (native language of the researcher and the interviewees). The interviews were guided by an interview guide of questions which were to be covered in each session. The 24 interviews were digitally voice recorded after the participants approval and then transcribed (for which an outside contractor was hired). Electronic copies of the recordings and the transcripts, as well of hard copies of the transcripts are stored for five years. Access to these is limited to the researcher only and password protected. The interviews varied between forty-five minutes and 2 hours.

The research project gained an Ethical Approval from the Research Ethics Committee of the Aston Business School of the Aston University. The Ethical Approval requirement was in compliance with the Research Ethic Guidelines and Research Governance document of Aston University. The identity of the four participating municipalities of this research study will be revealed in this thesis and subsequent articles. However, the confidentiality of personal information and the anonymity of all interviewees involved in this research study will be preserved. All interviewees will remain anonymous in recordings, transcripts and subsequent research outputs (i.e. dissemination conferences and/or academic research papers for publication in internationally recognised peer reviewed academic journals).

3.6 Analysis of data

In the analysis phase some specific methods were used that support the validity and reliability of the research study. In line with Modell's reasoning (2009a), authenticity and plausibility are seen as central aspects of validation. These analysis methods offer a framework for interpreting the findings and facilitate the linking of the data to the research questions.

Writing up the analysis of the data two approaches were employed. To gain familiarity with the data and to accelerate cross-case analysis, a pattern of thought in every individual case was sought for by doing a within-case analysis. Within-case analyses help to cope with the often enormous volume of data, typically involving detailed case study write-ups for each site. To compare the parameters of each case and to find common thoughts among the cases a cross-case analysis was done. Miles and Huberman (1994, p. 173) identify some of the reasons why cross-case analysis is important; chiefly generalisability and the deepening of understanding and explanation. Eisenhardt (1989, pp. 540-541) identifies three ways in which cross-case patterns can be searched for: (1) by selecting categories and then looking for within-group category similarities coupled with inter-group differences; (2) by selecting pairs of cases and then listing the similarities and differences between each pair; and (3) by dividing the data by data source. The cross-case analysis for this study included a comparison around themes (i.e. categories) and by data source. A useful assistance to this exercise was found from the Nvivo 9 software for qualitative research.

3.7 Interviewees

The 24 semi-structured interviews were conducted with the head of the Building and Housing Departments and employees working in the different sections of the department. The interviewees were selected based on their moderate or high experience with performance measurement. Among the interviewees is one female, the other 23 are male. To give an insight in the position of the interviewees in the organisation a distinction has been made about the decision making level and control activities of the interviewees. Managerial decision making and control activities can be categorised into three major types, namely strategic planning, management control and operational control (Emmanuel *et al.*, 1990, p. 96; Anthony and Govindarajan, 2006). Strategic planning is defined as being concerned with the setting and changing of overall strategies and objectives; management control involves monitoring activities and taking action to assure that resources are being effectively and efficiently used in accomplishing organisational objectives; operational control is concerned with carrying out specific tasks on a day-to-day basis. Insight in the position of the interviewees is given in table 3.2.

Table 3.2 Classification of 24 semi-structured interviews

	Breda	**Eindhoven**	**Emmen**	**Groningen**
Strategic	3	0	1	1
Managerial	3	3	4	3
Operational	3	1	1	1
Total (24)	9	4	6	5

Source: by author.

In three out of the four case municipalities the alderman, responsible for building and housing, was interviewed as well (i.e. strategic level). In Eindhoven this appeared not possible due to unrest at the political level during the data collection period and the resignation of the alderman. A complete overview of the interviewees can be found in appendix 4.

Because of the richness and depth of the evidence required, a multiple-case study has been selected for this research study. The following chapter describes the empirical context of the research. The intention of the empirical chapter is to give an idea about what will be studied and what not.

4. Empirical context

This chapter describes the empirical context of the research. The empirical context is important for understanding the development and key function of benchmarking within the country, the sector, the municipalities and the building and housing departments in particular. To start with, an overview is given of the institutional and political background of the Netherlands. In the subsequent section (4.2) a historical overview is given of performance management approaches in the Netherlands and reference is made to previous Dutch research in the field of PMM and municipal benchmarking. In section 4.3 an introduction is given to the building and housing sector, including an overview of the professionalisation towards quality improvement between 1999 and 2012. In section 4.4 an introduction is given to the benchmark Building and Housing Supervision (BWT), covering its history, philosophy and process steps. Section 4.5 gives insight in the setting of the four case municipalities and their organisational structure. Appendix 5 supports this chapter including detailed information about the benchmarking results for the year 2009.

4.1 Institutional and political background of the Netherlands

The Netherlands (16.5 million inhabitants) is a constitutional monarchy and organised as a decentralised unitary state. Municipalities form the lowest tier in government, after the central government and the provinces. As of January 2011 the Netherlands is divided into 418 municipalities and 12 provinces. Each of the sectors of government (national, provincial, municipal government and water boards) has its own responsibilities, with the central government providing unity through legislation and supervision. These work agreements ensure that municipalities and provinces possess a degree of autonomy (VNG, 2008).

The municipal council is elected every four years and is formally the highest authority in a municipality. The mayor is appointed by the crown and is the chairperson of the municipal council and also the chairperson of the municipal executive. The municipal council typically consists of groups or factions of elected representatives of political parties. The municipal council creates committees of members to monitor the various policy areas. The alderman responsible for a particular area of interest must give account of his policy position to the committee and the municipal council in general. Aldermen have a more political profile than the mayor and tend to be politically bound to a program (VNG, 2008). The governance of the municipalities changed from a monistic to a dualistic model in 2002. A basic feature of the dualistic model is that the distinction between the responsibilities of the municipal council and the executive committee are made more explicit. Since the implementation of this model in the Netherlands, the executive has been responsible for policy making and implementation, whereas the responsibility for boundary setting and monitoring the activities of the executive rests with the council (Tillema and Ter Bogt, 2008, p. 4).

The total budget of all Dutch municipalities (January 2011), amounts to almost 54 billion Euros (www.statline.cbs.nl). In general, the financing is organised in such a way that for the execution of tasks resulting from national policy, municipalities are compensated for the majority of the costs by contributions from national funds. Two major forms can be distinguished here: ear-marked funds and the general grant. In addition to these national funds (80%) municipalities have additional incomes (20%). These are local taxes (around 10% of the total income at the municipal level), charges (e.g., sewerage charges), rates (e.g., parking rates) and fees (e.g., construction permits). Municipalities have a major role in physical planning, public housing, transport, environment, social services, education, culture and welfare. Table 4.1 below gives insight in the categorisation of municipalities per number of inhabitants in the year 2011.

Table 4.1 Categorisation of municipalities per number of inhabitants (2011)

Number of inhabitants	Number of municipalities (418)	Percentage of municipalities (%)
< 5,000	6	1%
5,000 – 10,000	33	8%
10,000 – 20,000	117	28%
20,000 – 50,000	191	46%
50,000 -100,000	46	11%
100,000 – 150,000	12	3%
150,000 – 250,000	9	2%
250,000 >	4	1%
Total	418	100%

Source: adapted and modified by the author from the Central Office of Statistics

From the table above, it can be observed that municipalities in the Netherlands are relatively small. The relatively small size of the municipalities is one of the reasons why municipalities have entered into all kinds of co-operative ventures with neighbouring municipalities over the last decade. It is difficult for municipalities, and particularly the smaller ones, to organise all the tasks by themselves.

4.2 Performance management and benchmarking in the Netherlands

The past three decades have seen a surge of interest in the transfer of private sector management practices to the public sector. Following this international trend, Dutch municipalities have introduced some NPM-like changes in the 1980s and 1990s such as the adoption of output oriented planning and control and private sector management approaches. In an output budget (and the related planning and control documents) the budget functions are characterised by performance indicators. These changes were stimulated by the Public Management Initiative (PMI) as initiated by the Ministry of Home Affairs in 1988.

4.2.1 Historical overview

During the 1990s, more and more politicians and professional managers became interested in improving the performance of government organisations through various further initiatives. For example, several municipalities introduced further changes in organisational structures to make them more flat and/or to organise them around similar types of processes. Other popular initiatives involved internal contracting, the Balanced Scorecard and quality and performance management models, including the INK model, which is a Dutch equivalent of the international EFQM model (Ter Bogt, 2008a, p. 35). The table below gives insight in the historical developments of performance management approaches in the Netherlands.

Table 4.2 Historical overview of PM in the Netherlands

When	What	How
1979	Change in the Governments Accounts Act	The Act obliged the Dutch provinces to introduce accrual accounting from 1982
1982	Accrual accounting	Obligation for *provinces* to introduce accrual accounting
1985	Accrual accounting	Obligation for *municipalities* to introduce accrual accounting Some large Dutch municipalities changed control systems in order to facilitate a more business-like approach (e.g., Delft, Enschede, Tilburg and Groningen). Contract management was introduced and the focus on citizens and outputs strengthened.
1988	Start of Public Management Initiative (PMI)	The project started in 1988 and ended in 1995. The PMI project stimulated municipalities to apply private sector tools, such as output budgeting, responsibility accounting and cost allocation. Soon after the start, the PMI project was handed over to the municipal level, i.e. the Association of Netherlands Municipalities (VNG).
1990s	Focus on performance of government organisations	For example: further changes in organisational structures, internal contracting, the Balanced Scorecard and quality and performance management models such as the INK model (a Dutch equivalent of the international EFQM model).
Late 1990s	Intensified focus on performance of	For example: renewed focus on strategy, e-government, organisational culture, competences, and 'personal

	government organisations	development plans' for officials. Further, there have been new initiatives relating to 'integral management', contract management, benchmarking, and cooperation with other organisations.
1999	VBTB initiative	VBTB (from policy budget to policy accountability) – introduction of an outcome-based budget structure.
2002	The governance of the municipalities changed from a monistic to a dualistic model	A basic feature of the dualistic model is that the distinction between the responsibilities of the municipal council and the executive committee are made more explicit.
2002	Outcome budgets	Obligation for *municipalities* to introduce outcome budgets (i.e. program budgets). Program budgets, in addition to output budgets, should indicate the most important goals of policy programs as well as performances to be achieved.
2003	Outcome budgets	Obligation for *provinces* to introduce outcome budgets
2004 and beyond	Changes in financial management introduced	Local government organisations have taken initiatives to introduce changes in their financial management. For example: decentralised units started to present their own budgets and annual accounts and new financial administrative systems that make it easier to consolidate financial information from decentralised units were introduced.

Sources: Ter Bogt (2004); De Bruijn and Van Helden (2006); Ter Bogt (2008a and 2008b)

As can be observed from the table, since the late 1990s many organisations have increased or renewed their focus on strategy, e-government, organisational culture, competences, and 'personal development plans' for officials. Voluntary benchmarking is one of the imports from the late 1990s. Further, there have been new initiatives relating to 'integral management', contract management, and cooperation with other organisations. The series of reforms culminated in 1999 with the 'VBTB' initiative – an outcome-based budget structure. Since 2002, following the introduction of the so-called dualistic model in local government, the new Governments Accounts Act has required the Dutch provinces and municipalities to make a distinction between policy-based outcome budgets (i.e. program budgets) and output budgets (i.e. product budgets) (Ter Bogt, 2008b, p. 212). More recently, local government organisations have taken initiatives to introduce changes in their financial management.

4.2.2 Previous studies into PMM and benchmarking

Several publications can be found discussing the purpose and use of performance information in the Dutch public sector. The implementation of a change process based on the availability of performance information has been studied by Ter Bogt (2008a). The empirical findings of his study are based on 23 interviews with politicians and professional managers in twelve municipalities and two provinces. Most of the interviewees in the study of Ter Bogt were of the opinion that all in all the various changes had a slightly positive effect on the functioning and effectiveness of their organisations. However, they said that there was little or no information, either qualitative or quantitative, about the efficiency and effectiveness of their organisations, or about the effects of PMI and related management changes on efficiency and effectiveness. In addition to this the interviewees were of the opinion that the positive effects of management changes were not mainly due to 'technical' reforms like changes in organisational structure and planning and control systems, but rather to 'softer' changes such as increased attention to competences, the 'attitude' of employees and organisational culture. In another article published in the same year (2008b) Ter Bogt concludes that since 1985 the accounting changes, even though they were not a success in a technical sense, did bring about some effects in organisational culture and organisational learning that are in keeping with the ideas of NPM.

Both Ter Bogt (2004) and Jansen (2008) focus on the use of performance information by politicians. Based on a survey among Dutch aldermen Ter Bogt observes that many aldermen see little value in the output-oriented performance information that is available in the planning and control documents of their organisations. In general, aldermen seem to prefer rich, verbal information to sources of written information. Jansen adds that for politicians, there needs to be an incentive to use information about internal processes and outputs. He observes that politicians have a different perspective on performance, as compared to the internal process perspective and output perspective which are implied in NPM. Managers have an internal perspective, whereas politicians have a citizen perspective and a financial perspective on performance.

De Bruijn and Van Helden (2006) researched the factors that explain the success or lack of success of three examples of performance measurement in the Dutch public sector; namely performance measurement for planning and control in municipalities, benchmarking of waste water treatment by water boards, and performance measurement of educational services. The authors conclude that performance measurement systems in the public sector will be successful if these systems are developed and used in an interactive way between managers and professionals. Referring to Simons (1995), interactive systems focus on communicating and implementing the organisation's strategy. The purpose of an interactive system is to promote debate related to the assumptions underlying the organisation's strategy and ultimately to promote learning and growth. De Bruijn and Van Helden (2006) conclude that without an interactive system, systems of performance management can create strong incentives for perverse behaviour. In addition to this, the authors observe that in a consensus-based culture as in the Netherlands, the idea of performance management being embedded in the behavioural model might develop more easily than in a hierarchical culture. The authors observe a tension between the feminine world of consensus and the masculine world of performance control and refer to the risk that without a behavioural-oriented approach, systems of performance management can create strong incentives for perverse behaviour. Teelken (2008) studied the difficulties experienced with implementing performance measurement in the Dutch higher education and healthcare sector applying a longitudinal and cross-sectional comparison. The author observes that the implementation of PM systems is slower than intended and seems to occur outside the primary process of the organisation. The 'human' side of PMS should be able to acknowledge and bridge the gap between the primary process and the organisation of performance measurement.

Many interviewees included in the study of Ter Bogt (2008a) regarded benchmarking as a potentially important instrument in the future. They hoped that carefully considered benchmarking would enable them to obtain indications of the relative performance and efficiency of their organisations. Moreover, they thought that benchmarking could help them increase their organisations' transparency and accountability, and thereby help them to take account of citizens' desires. In the last decades several studies can be found related to public sector benchmarking in the Netherlands. Van Helden and Tillema (2005) developed a theory with respect to the various response patterns of public sector organisations to benchmarking, and supported this theory by empirical research into a benchmarking project of waste-water treatment by Dutch water boards. The theoretical framework developed by the authors is based on a combination of institutional and economic reasoning. In a later study among Dutch Water Boards, Tillema (2007) concludes that public sector organisations may ignore information (including benchmarking information) that indicates that their relative performance is poor. In addition she concludes that a benchmarking culture and having powerful managers is of great importance. Braadbaart (2007) studied if voluntary benchmarking indeed boosted the performance of public sector organisations in the Netherlands water supply industry. The author applied a quasi-experimental method to 1989-2000 time series data on benchmarking and non-benchmarking water utilities in the Netherlands. He came to the conclusion that benchmarking immediately enhanced transparency, but only affected economic performance after benchmarking information entered the public domain. In his view this confirms that benchmarking enhances transparency and performance.

In the Netherlands, benchmarking is seen as a potentially important instrument in the future and indications exist that benchmarking enhances transparency and performance. At the same time it can be observed that key issues identified in previous studies into PMM and benchmarking in the Netherlands are in line with the main themes that emerged from the review of the articles as presented in chapter 2, such as the role of inner and outer contextual factors in measurement system development.

4.3 Building and Housing Sector

The administrative layers in the Netherlands have different responsibilities and powers with respect to building and housing. In many countries the guaranteeing of the safety, health, and durability of the constructed environment is a traditional government task. In the Netherlands the Ministry of the Interior sets the framework whereas planning and implementation of these tasks has to be worked out at the municipal level. The provinces have fewer responsibilities with respect to building and housing, but they take care of the coordination between municipalities and may formulate a building and housing mission statement at the provincial level.

4.3.1 Organisational structure and tasks

A typical structure of the Building and Housing Department includes permits and inspection sections. The permit section is responsible for the intake of the permit request and the accounting or exemption of the permit. The section checks whether the submitted design meets the local zoning requirements and whether the construction meets the technical and environmental requirements on safety, durability, and health. Additionally, an independent committee of professionals (aesthetics committee) judges whether the design fits into the existing buildings and landscape from an aesthetical point of view. A permit will only be granted if the proposed design passes all three tests (Hoogwout and Te Velde, 2004, p. 234). The statutory rules can be found in the Housing Law, the Building Act and the Building Code. The Law on Spatial Planning provides urban and planning requirements. The permit to be applied for depends on the type of structure. According to the Housing Law, building permits can be divided into three categories: unlicensed, light license and regular license. Per the 1st of October 2010 new regulations have been introduced for the issuing of building permits. By the introduction of the so called all-in-one permit for physical aspects (WABO), 26 permits and 1600 forms have been integrated in one permit. The technical rules still apply, but the procedures to obtain a permit changed drastically. In 2010, a total of 25,141 building permits were issued in the 418 municipalities for both new buildings (49%) and remaining work (51%). From the total permits issued, 80% of the value applied to new buildings and 20% to remaining work (www.statline.cbs.nl). The inspection section is responsible for the supervision and control of licenses granted and the monitoring and enforcement of construction and land use laws and regulations.

Since the beginning of the 21st century the objective of successive governments at the national level has been to annually construct 100,000 new homes to reduce the housing shortage. However, year after year this objective has not been achieved. On average, the housing stock in the period 2000-2009 increased by 72,000 dwellings per year, including 54,000 in the occupied sector and 18,000 in the rented sector. Moreover, due to the credit crisis, the number of completed homes fell sharply to a level of 56,000 dwellings in 2010. Local governments saw their income decline in 2009 due to disappointing land sales and income from building fees. Dutch municipalities suffered a loss of Euro 414 million in 2009 on land (against a profit of Euro 600 million in 2007 and 2008) (CPB, 2011, pp. 143-150). In 2010, the housing stock in the Netherlands included a total of 7,217,803 houses (a unit that is assigned for living by a private household).

Around 2000 some disasters and serious accidents occurred in the Netherlands where buildings have collapsed or otherwise were found not to meet the requirements. Laws and regulations were also not always properly maintained or lacked the necessary supervision. These unfortunate events led to the professionalisation of enforcement and licensing on a national level. The building and housing sector received national attention and as a consequence, the interest in the benchmark BWT (introduced in 1999) intensified. More recently, the building and housing sector has operated in a turbulent context. The developments in environmental law, the introduction of the WABO and the debate surrounding the introduction of the regional implementation services (municipalities and provinces will cooperate to improve the quality of the licensing, supervision and control of complex environmental related implementation tasks) have occupied the minds considerably. In addition to this, financial cuts also provide a tension between achieving social outcomes and the retreating state.

4.3.2 Quality improvement

In table 4.3 a brief historical overview is given of the professionalisation towards quality improvement of the building and housing sector between 1999 and 2012. Information about changes in regulation and quality standards is included in this table, since these are seen as relevant for the professionalisation towards quality improvement and affects the content of the benchmark BWT.

Table 4.3 Historical overview of the professionalisation towards quality improvement of the building and housing sector between 1999 and 2012

What	When	Characterisation
Benchmark BWT	1999	Launch of the benchmark BWT by PricewaterhouseCoopers (PwC) Consulting. Since 2003, the benchmark has been managed by the consulting company Zenc.
Platform BWT Large Municipalities	1999	This platform for municipalities with >100.000 inhabitants has laid the groundwork for a partnership approach. In fact, the instrument is a combination of existing methods of comparison with a long-term intention for co-operation. The 13 municipalities that participated both in the platform and the benchmark in 1999 included: Almere, Arnhem, Breda, Den Haag, Eindhoven, Emmen, Enschede, Heerlen, Helmond, Maastricht, Utrecht, Venlo en Zwolle (Breda, Eindhoven and Emmen are case organisations in this research study).
Association BWT	2003	The Association BWT is replacing the Platform BWT Large Municipalities and has grown into the leading association in the sector. The association was initially created for the sole purpose of the management of the "collective quality standards building permits" system (LTP – previously called CKB). These quality standards are taken as a reference point in the benchmark.
Launch of the knowledge base for the building and housing sector (www.bwtinfo.nl)	2005	An online expert system that contains the domain knowledge of the sector. The objective of the system is to improve the mutual knowledge exchange at the municipal level for a more professional, effective and efficient realisation of the building and housing sector.
Introduction of certification of the benchmark BWT by the Association of Netherlands Municipalities (VNG)	2005	The benchmark BWT received a certificate in 2006 (the first round of certificates awarded by VNG), 2007, 2008, 2009 and 2010.
Launch of the national key protocol for the collective quality standards for building permits (LTP – previously called CKB)	2006	This is a national method for testing building plan applications to the structural laws and regulations. The key protocol LTP supports quality thinking and provides for recording of the actual test results. These quality standards are taken as a reference point in the benchmark.
Report Commission Dekker	May 2008	The commission proposed to reform the building regulations and building code structure.
Report Commission Mans	July 2008	Research into the enforcement of the regulations in the field of environmental law. In the proposed model, municipalities retain their formal enforcement powers, but make use of the regional implementation services (RUD). The commission proposes to establish regional implementation services in the country.

What	When	Characterisation
Report Commission Oosting	May 2009	The commission proposed a reorganisation of the inter-administrative supervision arrangements in the sector.
Launch of an approach to construction incidents for structural safety (www.abcmeldpunt.nl)	October 2009	ABC is an initiative of the Platform Structural Safety. The registration system includes the registry of construction errors that relate to structural safety. The goal of this registration is to learn from past mistakes and hence ensure safer buildings. Alerts can be done by persons working in the sector. After that the alerts will be made anonymous and entered into the database. By means of analysis causes and lessons are determined.
Platform BWT Large Municipalities	October 2009	Reestablishment of the Platform BWT Large Municipalities (>100.000 inhabitants + Leeuwarden, Helmond and Venlo)
Final report quality standard development – version 2.0 (KPMG)	December 2009	The aim of this project is to develop quality criteria for the authorisation of the all-in-one permit for physical aspects (WABO) and the monitoring and enforcement of legislation for the physical environment.
All-in-one permit for physical aspects (WABO)	1st October 2010	The Dutch act 'WABO' that lays down the rules for granting an all-in-one permit for physical aspects came into force on 1st of October 2010. The new Act has replaced around 25 separate permits. The introduction of the permit involves substantial changes for permit-issuing authorities. The Act as well regulates coordination. Government authorities involved in a particular application are required to cooperate with each other to take one harmonised decision, issued by one competent authority.
Digitisation of customer contact	1st October 2010	This initiative is part of the introduction of the all-in-one permit for physical aspects (a result of a pilot study implemented by the Association BWT).
Launch of the WABO Benchmark	2011	Successor of the benchmark BWT that was launched in 1999. The WABO benchmark includes the aspects that are related to the all-in-one permit for physical aspects. The WABO benchmark received the quality certificate in 2011 from KING.
Program Execution with Ambition (PUmA)	2011	The organisation and execution of licensing, supervision and enforcement of the regulations in the field of environmental, building and space has structural problems. To solve these, the "Program Execution with Ambition" (PUmA) has been designed. Eventually this will lead to the formation of the regional implementation units. The project Quality (component of PumA) contributes to increasing the quality of implementation of tasks in the building and housing supervision field. A self-assessment tool has been developed as part of the project. With this tool, the municipalities can assess how the organisation

What	When	Characterisation
		compares to the reference set of quality standards as developed by KPMG in 2009. The central premise of the tool is self-monitoring of the quality and the promotion of quality thinking within the organisation. However, the criteria also serve as a tool for determining if services need to be handed over to the regional implementation service by January 2013. Between 30 and 32 regional implementation services (RUD) will be established in the country as a component of PUmA. Municipalities and provinces will cooperate to improve the quality of the licensing, supervision and control of complex environmental related implementation tasks.
Certified building plan reviews (www.bpt.nu)	November 2011	It is possible for municipalities to outsource the building plan review against the Building Act by a certified organisation. The certified building plan review aims to offer municipalities and clients an objective and clear understanding of the test results.
Update of quality standard development – version 3.0 (on-going)	2012	Project to update the quality criteria for the authorisation of the all-in-one permit for physical aspects (WABO) and the monitoring and enforcement of legislation for the physical environment.

Source: author, based on several internet sources.

The changes in regulation and quality standards as included in the table indicate a profound development towards quality improvement in the building and housing sector over the last decades. A couple of observations can be made. Firstly, the role of the association (previously platform large municipalities) can be seen as significant in both developing the quality standards and in building a knowledge base in the building and housing sector. A second observation is that a continuous search exists for defining quality in the sector. The national key protocol for the collective quality standards for building permits (LTP – previously called CKB) supports quality thinking and provides for recording of the actual test results. Quality standard development received a new impetus in 2009 with the development of quality criteria for the authorisation of the all-in-one permit for physical aspects (WABO) and the monitoring and enforcement of legislation for the physical environment. In 2011, a new program (PUmA) acknowledges the importance of the inner context of the municipalities in the quality discussion. As a consequence, a self-assessment tool has been developed. The central premise of the tool is self-monitoring of the quality and the promotion of quality thinking within the organisation. A third observation is that the quality of the benchmark tool itself is addressed as well with the introduction of certification of the benchmark BWT by the Association of Netherlands Municipalities (VNG). Apart from offering a historical overview of the professionalisation towards quality improvement, table 4.3 gives insight in the increasing attention and search for quality improvement in the building and housing sector and the move towards a more inclusive approach by focusing on both the quality of achievement and the quality of actions in the sector.

4.4 Benchmark Building and Housing Supervision

Voluntary benchmarks in the Netherlands are designed and offered by a sector representative organisation or consultancy firm. Municipalities pay an annual fee to participate in the specific benchmark cycle, of which the outcomes are of a confidential nature. The benchmark BWT has been selected as the basis for this research study. Important arguments were: maturity, expected access to data, and the amount of data already available. The selected benchmark is meeting the quality criteria, for which it obtained a quality mark, as formulated by the Association of Netherlands Municipalities (VNG) and the Quality Institute for Netherlands Municipalities (KING). With reference to the benchmarking typologies according to their uses as presented in table 1.5, the benchmark BWT can be classified as a voluntary benchmark, promoting a voluntary and collaborative learning process. The term benchlearning as introduced by Karlöf and Östblom (1995) will be used to refer to learning from the benchmark BWT (see as well section 1.2.1).

The benchmark BWT was started in 1999 on the initiative of a platform for larger municipalities (> 100,000 inhabitants). In this platform 35 municipalities worked jointly on projects such as the project Collective Quality Norms for Building Permits (CKB, later replaced by LTB). The platform transferred in 2003 into the Association for Building and Housing Supervision for professionals in the field of building and housing. The association continued the benchmarking initiative as started by the platform in 1999 on a yearly basis. The consultancy firm that carried out the benchmark in 1999 (PwC Consulting) applied the so-called principle of co-creation; meaning that potential users of the results were invited to actively participate in the teams that identified, defined and fine-tuned the performance indicators. Since 2003, the benchmark has been implemented by a Dutch consultancy firm named Zenc (also referred to as the benchmark provider). The performance indicators included in the benchmark are evaluated on a yearly basis jointly with the potential users.

The questionnaire of the benchmark BWT is constructed around three perspectives: the client perspective (quality of service delivery); the management perspective (effectiveness) and the organisation perspective (efficiency). The *quality* of the service delivery can be seen as the perspective of how the clients see the licensing process. The applicant benefits from a quick service delivery, a careful balancing of the interests, a minimisation of the administrative fee and a client focused and accessible municipality. The *effectiveness* of the licensing process can be seen as the management perspective. The interest of the management is that the stated policy goals with regard to building and housing supervision are achieved. The effectiveness is measured, inter alia, in achieving policy goals. Also, the extent to which the municipality is able to meet its legal duties in practice is important. The *efficiency* is the perspective of the organisation, and includes looking at the number of applications processed per fulltime-equivalent, the cost of a license and the percentage of refused and declared inadmissible applications. The benchmark is based on the philosophy that all three perspectives need to be in balance with each other. A summary of the three perspectives and key performance indicators for each perspective is given in appendix 5, including the figures for the case municipalities included in this research study for the year 2009 (data 2008).

Around 25 municipalities participate in each benchmarking cycle. As part of the cycle at least two and sometimes three meetings are organised for participating municipalities (the learning network). The composition of the learning network changes each year, since not all municipalities participate annually. Different municipalities participate in each cycle, changing the composition of the learning network. The benchmark provider acts as the project secretariat, guaranteeing the quality in the process. A first meeting of the learning network takes place to evaluate the performance indicators proposed to be included in the benchmark. If necessary, performance indicators will be fine-tuned in this stage. The collected performance information is made available to the participating municipalities on an individual basis in a tailor-made report. The individual scores of the municipality are compared with the average, the highest and the lowest score of the benchmark. If available, the scores of the municipality will be compared to results in previous rounds. A second learning network meeting is scheduled to discuss the aggregated data from all municipalities participating in the benchmark cycle. In this meeting, outcomes are assessed and discussed. It is also examined whether there is an explanation for striking deviations from the benchmark averages. If requested by the participating municipalities a third meeting will be organised by the benchmark provider to discuss specific themes in more depth.

4.5 Case organisations

In this section insight is given into the setting of the case municipalities and their structure. The selected case municipalities are: Breda, Eindhoven, Emmen and Groningen. The table below gives insight in some key statistical figures of the four case municipalities for the year 2010.

Table 4.4 Key statistical figures of the case municipalities for the year 2010

2010	**Population**	**Population density per km²**	**Housing stock**	**Housing density per km²**	**Income building permits**	**Urbanisation level**
Breda	173,299	1,368	76,284	602	4,846,000	2
Eindhoven	213,809	2,438	96,993	1,106	7,453,000	2
Emmen	109,491	325	46,800	139	2,627,000	4
Groningen	187,298	2,393	85,117	1,087	7,568,000	1

Sources: Compiled by the author based on CBS (2010) and www.statline.cbs.nl

The housing stock indicates all residential buildings intended for permanent habitation by a private household. The urbanisation level is a categorical classification of municipalities based on the area address density (the average number of addresses per square kilometre). Five groups are distinguished: (1) very high urbanised municipalities: area address density of 2,500 addresses or more per km²; (2) highly urbanised municipalities: area address density of 1,500 to 2,500 addresses per km²; (3) moderate urbanised municipalities: area address density of 1,000 to 1,500 addresses per km²; (4) little urbanised municipalities: area address density of 500 to 1,000 addresses per km²; and (5) non-urban municipalities: area address density of less than 500 addresses per km².

Breda is a city situated in the South-West of the Netherlands, classified as a highly urbanised municipality. The municipality of Breda is a member of several national platforms (e.g., the Association of Municipal Secretaries and the network of municipalities with more than 100,000 inhabitants). The municipality of Breda went through several reorganisation processes in the last decade. In the year 2006 the Building and Housing Department was reorganised. The department became part of the Public Services Division at the end of 2006. With this the organisation anticipated the arrival of the all-in-one permit for physical aspects. At that time the department was divided in the sections Permits and Inspection. The staff members were divided in three regional divisions: East, West and Centre and were all headed by a manager with a multi-disciplinary team. At the time of the data collection for this research study (2009-2010), Breda's civil service consisted of 10 sectors and two staff departments and could be typified as a matrix organisation. The matrix structure combines both functional divisions and product groups. A cultural change process was introduced in 2009 and discussed within the department early 2010. The approach (appreciative enquiry) is a tool to set innovations in motion in the form of interviews, meetings and gatherings, where involvement and process are more important than the result. In 2010, the housing stock in Breda included a total of 76,284 houses of which around 60% are homes for sale. The total revenue of issuing building permits in 2010 is Euro 4,846,000. The previous period (2005-2006) yielded fewer houses than desired (caused by a delay in one of the neighbourhoods and the economic downturn). From the figures it can be concluded that the average revenue of the building permit per residential building is Euro 64. The municipality participated in the benchmark BWT in the years 1999, 2002, 2005, 2007, 2008 and 2009.

Eindhoven is a city situated in the southern part of the Netherlands, classified as a highly urbanised municipality. The municipality of Eindhoven went through several reorganisation processes over the past years. A reorganisation in 2008 resulted in the putting together of several departments (permit, supervision and enforcement) to improve the internal communication between the departments. In 2009, the Building and Housing Department was divided into the following two sections: Permits and Inspection. The motivation for the most recent reorganisation (2009) can be found in the intended introduction of multiple permits (WABO) in 2011. In 2010 the housing stock in Eindhoven included a total of 96,993 houses. The division between rental housing and ownership is about 50/50. Eindhoven strived for an average production of 1,500 new dwellings per year in the period 2005-2010. The total revenue of issuing building permits in 2010 is Euro 7,453,000. From the figures it can be concluded that the average revenue of the building permit per residential building is Euro 77. The municipality participated in the benchmark BWT in the years 1999, 2002, 2005, 2006, 2008 and 2009.

Emmen is a city situated in the northern part of the Netherlands, classified as a little urbanised municipality. The municipality of Emmen is the second largest municipality in the Northern Netherlands and offers employment to about 1,100 civil servants. The municipal territory covers no less than 35,000 hectares and the city still has an abundance of land available for residential development and business establishment. The municipality went through a reorganisation process in 2008 focusing on cost-efficiency, and introduced among others the concept of flexible working. At the time of the data collection for this research study (2009-2010), Emmen had an administrative organisation including six departments, the executive administrative unit, and the fire brigade. In 2010, the housing stock in Emmen included a total of 46,800 houses. The total revenue of issuing building permits in 2010 was Euro 2,627,000. From the figures it can be concluded that the average revenue of the building permit per residential building is Euro 56. The municipality participated in the benchmark BWT in the years 2005, 2007, 2008 and 2009.

Groningen is a city situated in the most northern part of the Netherlands, classified as a very high urbanised municipality. Groningen offers employment to about 3,548 civil servants in December 2010 (a total of 3,130 FTE's). The number of staff is high compared to many other medium-sized cities. This is because of the caring function of the municipality for the environment (for example youth, health and theatre). Groningen's civil service consists of six sectors and two staff departments. In 2009, the Building and Housing Department was divided in the following three sections: Inspection, Policy and Support, and Housing and Building Desk. In 2010, the housing stock in Groningen included a total of 85,117 houses. The total revenue of issuing building permits in 2010 was Euro 7,568,000. From the figures it can be concluded that the average revenue of the building permit per residential building is Euro 89. The municipality participated in the benchmark BWT in the years 2005, 2008 and 2009.

In the subsequent chapters (5 to 7) the findings of the empirical phase are brought together with the insights drawn from the literature review.

5. Municipal benchmarking and organisational learning

"Interaction can develop shared meaning and perspectives, which is the basis of knowledge. This perspective contends that knowledge is developed through interaction and within practice, and any attempts to move it from the context of this interaction will be problematic".

(Rashman *et al*., 2009, p. 477)

5.1 Introduction

The literature on municipal benchmarking to date is not clear about *how* a municipal benchmark shapes organisational learning within municipalities and whether and how the "human" side effect the impact of municipal benchmarking on organisational learning and performance. Pettigrew refers to this as the process of change (Pettigrew, 1987, p. 658): *"The process of change refers to the actions, reactions, and interactions from the various interested parties as they seek to move the organisation from its present to its future state"*.

Bringing organisational learning into research on municipal benchmarking should provide a better foundation for understanding why some municipalities are better able to learn and why only some of the useful performance information is embraced. The expanded 4i framework of organisational learning (see introduction in chapter 1, section 1.3.2) is used here to support the analysis. To have a better insight in how levels of analysis and learning processes in municipal benchmarking affect benchlearning, the following research question and related sub-questions are explored in this chapter:

Which levels of analysis, learning processes, learning barriers and supporting factors can be observed in municipal benchmarking in the building and housing sector in the Netherlands? This research question is further divided into two sub-questions:

(a) In what way can the 4i framework of organisational learning be adjusted to provide a helpful starting point for studying municipal benchmarking?

(b) What empirical observations can be made when applying the adjusted learning framework to the four case studies?

This chapter examines the above observation on unexamined causal links using insights based on the 24 semi-structured interviews conducted between November 2009 and September 2010 in four Dutch municipalities. To analyse the data two approaches were employed. To gain familiarity with the data and to accelerate cross-case analysis, a pattern of thought in each individual case was sought for by doing a within-case analysis. To compare the parameters of each case and to find common thoughts among the cases a cross-case analysis was done.

Section 5.2 provides an introduction to the organisational learning literature as a follow up to the conclusion at the end of chapter 2 (literature review) that the learning potential of a benchmark is an important issue. The studies into the learning potential of a benchmark may be strengthened when insight is given in how a municipal benchmark shapes organisational learning within local government agencies. To this end, this section puts specific emphasis on the relation between "organisational performance" and "organisational learning", including some additional literature. The subsequent sections (5.3 to 5.6) present the empirical data for the research question. In section 5.3 the levels of analysis and related learning processes that can be observed in municipal benchmarking are presented and analysed per case study. Section 5.4 and 5.5 respectively give more insight in the learning barriers and supporting factors between the levels using the learning barriers' framework developed by Schilling and Kluge (2009). In the last section (5.6) some concluding observations are made.

5.2 Organisational learning

There has been considerable research into the key stages to carry out benchmarking (Moriarty and Smallman, 2009; Andersen *et al*., 2008; Anand and Kodali, 2008; Kyrö, 2003). However, movement through these stages involves more than sequential activities. Recent research on municipal benchmarking (e.g., Broadbent and Laughlin, 2009; Horton, 2006) and the literature on performance management in general give insight into how a municipal benchmark shapes organisational learning within municipalities by studying the "human" side of a municipal benchmarking system. This will allow for further understanding and explaining why people have different experiences with the incorporation and use of performance information. The literature suggests that municipal benchmarking is likely to deliver improved organisational performance when it is coupled to organisational learning.

First, the learning potential of a benchmark is an important issue. Moynihan (2005) observes that most results-based reforms target narrow process improvement (single-loop learning) rather than a broad understanding of policy choices and effectiveness (double-loop learning). Argyris and Schon (1978) argue that double-loop learning is necessary if practitioners and organisations are to make informed decisions in rapidly changing and often uncertain contexts. Their theory of action distinguishes the defensive, non-learning routines of "model one" behaviour from that of "model two" individual, collective and double-loop learning (Rashman *et al*., 2009, p. 469). Kouzmin *et al*. (1999) refer to the paradox that the outstanding pre-condition for learning in organisations is the creation of the "learning" organisation in structural and cultural terms. Senge (1990) has influenced both academics and practitioners, developing the notion of organisations as systems and popularizing the concept of the learning organisation (Rashman *et al*., 2009, p. 469).

Second, the studies on municipal benchmarking may be strengthened with a behavioural-oriented approach. Holloway *et al*. (1999) state, that benchmarking is only as effective as the people who apply it. The effective use of performance information derived from municipal benchmarking requires a major change in the attitudes of the people using the information.

Different theoretical conceptions of organisational learning have been developed (e.g., Easterby-Smith, 1997; Easterby-Smith *et al*., 1998; Kluge and Schilling, 2003). The theoretical conceptions that can be found in the literature emphasise the dual nature of learning as a process (perceiving and processing information, i.e. experience) and as a result (modified knowledge or skill). As the overarching goal of this chapter is to integrate existing evidence on the "human" side of municipal benchmarking, concentration is laid on the perspective of organisational learning as a process, rather than on its results. In line with Schilling and Kluge (2009, p. 338):

> *"Organisational learning is defined as an organisationally regulated collective learning process in which individual and group-based learning experiences concerning the improvement of organisational performance and/or goals are transferred into organisational routines, processes and structures, which in turn affect the future learning activities of the organisation's members".*

To have a better insight in how levels of analysis and learning processes in municipal benchmarking affect benchlearning, the 4i framework of organisational learning is used in this chapter to support the analysis of the four case municipalities. As presented in section 1.3.2, Crossan *et al*. (1999) developed the 4i framework of organisational learning, where learning processes of intuition-interpretation-integrating-institutionalisation interact with knowledge exploitation and exploration dynamics throughout three analysis levels: individual, group and organisational levels.

Intuition refers to a uniquely individual process which may happen within a group or organisational context. *Interpretation* is a personal explanation for one's self and for others about individual's insights requiring verbal manifestations and language development. *Integration* is a process through which shared understandings among individuals occur and coordinated actions through mutual adjustments result. Dialogue and conjoint action are crucial for developing shared understandings. *Institutionalisation* is a process of guaranteeing that routinised actions occur.

The central proposition of the 4i framework is that the 4i's are related in feed-forward and feedback processes across the levels. Through feed-forward processes, new ideas and actions flow from the individual to the group to the organisation levels. At the same time, what has already been learned feeds back from the organisation to the group and individual levels, affecting how people act and think.

Nearly every organisational learning process model describes interpretation as a single discrete phase. The learning episodes within four employment services agencies studied by Mausolff (2004) revealed, that every phase can be characterised by new information and new opportunities for interpretation of that information Mausolff (2004) categorised the general studies of organisational learning according to: (1) those emphasising the creation of stress in response to feedback; and (2) those describing the social construction of meaning. In the first category organisation members often did not take action unless performance was at crisis levels. However, the learning episodes studies by Mausolff (2004) revealed that stress was not a necessary condition for learning in the studied organisations. Different works in the second category all share the assumption that meaning creation is fundamentally a social process. In this research study it is assumed that meaning creation is fundamentally a social process taking place throughout different levels of analysis within the municipal organisation.

Theorisation on the process of integration was previously made by Grant (1996), who suggests mechanisms and characteristics of integration. The two main proposed mechanisms of integration were (1) directions, in which knowledge is captured and codified into rules, procedures and operating manuals; and (2) routines, in which more difficult to capture tacit knowledge is conveyed through sequential patterns of interaction (Grant, 1996). Van Winkelen and McKenzie (2007) explored the question "in what ways planned learning activities in organisations can be integrated more effectively" and discovered that: *"factors that positively influence integration were found to include widespread recognition of the business value of both individual and organisational learning, high level sponsorship that acts as a bridge across functional boundaries and line managers adopting an integrating approach to learning in managing their people and the tasks they undertake. Factors that negatively impact the adoption of an integrated approach were found to include the lack of mechanisms to coordinate across functions and a culture in which functional managers feel unable to change practices"* (p.1).

Wiseman (2008, pp. 53-55) explored the process through which organisational knowledge is institutionalised in an international mining company by use of the 4i framework of organisational learning. The study found that the process of institutionalisation involves a number of sub-processes in which knowledge is carried through the organisation through routines, symbolic systems and relational systems. The author explains that routines arise in areas where action is repeated to the point of eliminating individual thought in the enactment of the action. Symbolic systems can be seen as another sub-process, carrying institutional knowledge through culture, which encompasses the organisation's notions of rules, values, classifications, representations and logic. Institutions can also be carried and influenced by the patterns of expected behaviours that form the relationships found in networks of organisational positions.

The literature suggests that municipal benchmarking is likely to deliver improved organisational performance when it is coupled to organisational learning. At the same time, it can be observed that nearly every organisational learning process model describes interpretation as a single discrete phase. The subsequent section describes the analysis levels and related learning processes that can be observed in municipal benchmarking, structured along the 4i framework of organisational learning.

5.3 Levels of analysis and related learning processes

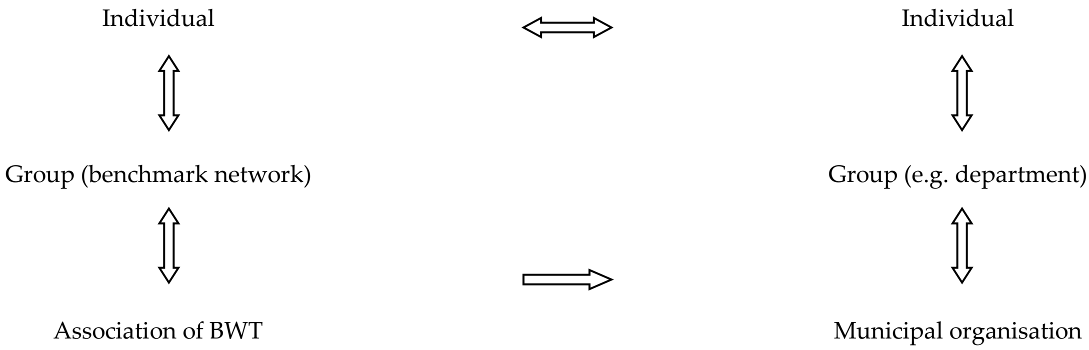
In municipal benchmarking, the benchmark network rather than the levels as indicated by the 4i model is often the initial locus of activity. The benchmark network that is offered as part of the Building and Housing Supervision benchmark is made up of individuals who do not necessarily work together regularly or form part of a professional community. Learning takes place external to the municipality and beyond everyday practice. Meier and O'Toole (2003, p. 690) define a network as "a pattern of two or more units, in which not all major components are encompassed within a single hierarchical array". The benchmarking network can be seen as a learning network, similar to learning networks within an organisation. This section describes the analysis levels and the related learning processes that can be observed in municipal benchmarking and gives insight into how these affect benchlearning in the four case studies.

5.3.1 Levels of analysis

Participants in the benchmark network are representatives of municipalities that have subscribed to take part in the annual benchmark cycle. Individual ideas develop within this benchmark network (inter-organisational group), and individuals (as representatives of their municipality) ultimately share those ideas with group members and the wider organisation at their own municipality. Municipal benchmarking is a process where performance comparisons drive the individual, group and organisational learning. An initial shared understanding is built within the benchmark network and thus external to the municipality. Rashman *et al*. (2009, p. 477) observe that this contends that any attempt to move it from the context of this interaction will be problematic. According to the authors it is important to describe the context-specific factors for a level of learning and to describe the nature of different participating organisations, as well as the network structure itself. The authors argue that the environment in which an organisation is operating can have a profound influence on the learning process. External factors may impact in specific ways upon the learning process and the organisational ability to mobilise knowledge (p. 478). Hodges and Grubnic for example observe (2010, p. 46) that an issue of importance in local government collaboration is likely to be the extent to which the partnership is driven by externally imposed performance indicators rather than enabling the partnership to develop its own internal mechanisms for performance and accountability.

Although the 4i framework provides a compelling basis for research on organisational learning, it neglects the inter-organisational level present in municipal benchmarking. Consequently, a wider conception of learning is suggested for benchmarking where learning processes occur in two parallel communities (network and municipality) with the individual as a linking pin (see figure 5.1). This conception assists in examining how context may shape learning drivers in the field of benchmarking and influence the learning process.

Figure 5.1 Levels of analyses in municipal benchmarking



Source: the author

In the case of municipal benchmarking, learning processes interact with knowledge exploitation and exploration dynamics throughout four analysis levels (i.e. inter-organisational, individual, group and organisational levels). The characteristics of each community (context) will influence benchlearning. The individual has a very important role, being the "bridge or linking pin" between the benchmark network and the organisational group within the municipal organisation itself. At the same time a feedback process can be observed from the Association of BWT towards the municipal organisation. An example is the involvement of the Association in the formulation of Collective Quality Norms for Building Permits that apply for all municipalities. This more complex view towards learning processes in municipal benchmarking is better served to support the analysis of municipal benchmarking than the linear 4i framework of organisational learning.

5.3.2 Learning processes

With municipal benchmarking, new ideas and actions based on inter-organisational performance comparisons flow from the benchmark network (inter-organisational level) to the context of the group and the organisation with the individual as a linking pin. At the inter-organisational level a process of interaction takes place between the members of the network, influencing the individual intuition and interpretation process. Interaction is the basis of simultaneous knowledge construction and transfer (Bate and Robert, 2002; Newell *et al*., 2003; Rashman *et al*., 2009). The interaction that takes place in the benchmark network, which serves to develop the understanding of the individual member of a municipality, also helps to identify perceptions and ideas in general. The ideas developed at the benchmark network require integration and institutionalisation at the level of the Association of BWT or within the municipality itself. It is recognised that there may be bottlenecks in the ability of the municipalities to absorb the feed forward of the learning from the benchmark network to the municipal organisation. Within the public sector, individuals may belong to a multitude of shifting groups and networks, many of which do not necessarily work as a team or operate as a community of practice (Rashman *et al*., 2009, p. 473). The individual as a representative of the organisational unit (Building and Housing Supervision) has a crucial role in transferring the knowledge from the benchmark network to his/her own group and the organisation, with the objective of integration and institutionalisation for improved performance.

5.3.3 Case studies

In this section empirical observations are made by applying the dynamic learning process to the four cases. The ability of the four cases to absorb the feed forward of the learning from the benchmark network, individuals, groups and the organisation is discussed by describing benchmark-related actions of each case organisation. Although a detailed investigation of organisational learning over time by Berends and Lammers (2010) showed that the sequence of activities that produced learning outcomes neither represented a single learning cycle nor a succession of multiple, discrete learning cycles, the levels of learning that can be observed in municipal benchmarking are taken as an analytical lens to structure the empirical data. For each case municipality an analysis is made of how processes of intuition-interpretation-integration-institutionalisation interact with knowledge exploitation (feedback) and exploration (feed forward) dynamics. As suggested by Rashman *et al*. (2009) the context-specific factors for a level of learning are described as well. The empirical observations relate primarily to the feedback process between the benchmark network and the organisational group within the municipal organisation itself. Although the feedback process from the Association of BWT towards the municipal organisation has not been the focus of the data collection, some observations can be made in section 5.3.4. The case analysis below will shed light on how municipal benchmarking leads to performance improvement. Similarities and differences between cases are listed.

Breda

In the municipality of Breda, several learning episodes of benchmarking could be detected from the interview excerpts:

- The benchmark results showed that there are several municipalities that have established policies for how to treat the existing regulations for the issuing of building permits. Because of this observation, derived from the performance comparison in the benchmark network, a working group has been established to formulate policy on this for the municipality of Breda as well;
- In parts, the municipality used the benchmark results for the further elaboration of the reorganisation (e.g., to analyse the workload and work pressure per department);
- The benchmark showed that the municipality of Breda was behind in turnaround times of the issuing of building permits in comparison with other municipalities. With the support of an external agency steps have been made for improvement;
- The benchmark showed a positive assessment of the personal contact that outside inspectors have with the client. This example has been followed by the plan reviewers, who now have more contact with the client as well;
- The benchmark showed that the deadline for the permits was exceeded more often than in other municipalities. This appeared to be a mistake in the registration

system because the system did not register the deadline. The system has been adapted;

- Following the results of the benchmark, a working group has been established to improve the quality of the permit review process.

The examples above show that dialogue and conjoint action have taken place in the municipality of Breda, leading towards shared understandings (integration). Based on the benchmark results some new actions have been proposed and embedded (institutionalisation), such as the formulation of new policy and structurally improved communication. So how was the knowledge transferred from the individual in the benchmark network to his/her own group and the organisation?

Respondents in the municipality of Breda referred to the benefit of several external network environments where personal explanations for one's self and for others about individual's insights in the building and housing sector (including the benchmark results) can be shared and discussed (interpretation). Within various expert groups verbal manifestations and language development occur to facilitate the interpretation process among the members. Reference, for example, is made to the expert group construction supervision, of which constructors are the participating members. Having access to the "source" of the information and experience of the members is seen as the key added value of these expert groups. Respondents lament the specialised fields where expert groups do not (anymore) exist. According to the respondents, these inter-organisational professional relationships have a stimulating function. The benchmark network of the Building and Housing Supervision benchmark is frequented by the head or deputy head of the department (i.e. managerial level). Interviewees are of the opinion that some content information can be obtained during the benchmark network meeting, but that not all issues and subjects come to the table since the participating members are not always knowledgeable in the subject.

Within the municipal organisation the head of the Building and Housing Department shares his individual interpretation based on the benchmarking report with the departmental management team with some room for interpretation from other team members. In the municipality of Breda, the report stays within the department and is not shared with the municipal council or the executive committee. According to the management team, benchmarking is a management tool. The head of the Building and Housing Department observes: "*Do we need to share with the municipal council the figures on sick leave of our department? That would be an administrative decision*".

The majority of the respondents at the tactical and operational level are of the impression that more could be done with the benchmarking results and that the management team (i.e. strategic level) is not paying sufficient attention to the discussion and dissemination of the results. Respondent 11 observes the following: "*I think that we can do more with the results than we actually do. Our own management does not take the time and effort to extract the figures and to discuss and disseminate these well*". Respondent 14 is offering some insight in the reasons for the perceived limited use of the benchmarking information: "*The benchmark is not part of our regular work. In first instance the report comes in a corner right. The question is, if those who receive the report find it interesting enough to do something with the results*". Respondent 16 is referring to a direction for improvement: "*We should actually make one person responsible to distil the learning points from the benchmark (…) which maybe leads to a joint exchange and formulation of action points. It is important to make one person responsible for the progress*".

It is remarkable that the respondents representing the tactical and operational level have the impression that little has been done with the benchmark information. At the same time several learning episodes of benchmarking could be detected of claimed uses and follow up actions from the interview excerpts. A possible explanation can be that the management team leads with action and uses a hybrid approach (a mixture of hierarchical and network strategies) for improvement: a decision has been taken at the management level and a working group is asked to start a process to discuss the future steps for improvement.

Another explanation can be that Breda is future-possibility oriented. Crossan *et al*. (1999, p. 526) refer to this as tacit knowledge: "*The expert no longer has to think consciously about* *action. Having been in the same, or similar, situations and recognising the pattern, the expert knows, almost spontaneously, what to do. Indeed, if asked to explain their actions, experts may be unable to do so".*

From the case description it can be observed that the municipality of Breda claims a wide variety of uses of benchmark information. It appears that the knowledge transferred from the individual in the benchmark network to his/her own group and the organisation results in the feed forward of the insights obtained. Nevertheless respondents at the tactical and operational level are of the impression that more could be done with the benchmarking results. It appears that professional expert groups further stimulate the learning process.

Eindhoven

In the municipality of Eindhoven, some learning episodes of benchmarking could be detected from the interview excerpts:

- Insight was given to the alderman into the level of building supervision that was offered in comparison with other municipalities. This comparison, on the basis of benchmark figures, offered sufficient explanation why particular choices had been made and how prioritisations were made within the department.
- It was observed from the performance comparison that the turnaround times of permit requests in another municipality (with whom Eindhoven frequently compares) were much longer, but at the same time they scored ten times higher on customer experience. Therefore more customer guidance in the processes was introduced in Eindhoven.
- The performance comparisons in the benchmark network showed that the building sites in Eindhoven were not visited very frequently in comparison with other municipalities. This performance information appeared inaccurate. After further analysis it appeared that the twelve persons working in the field did visit the building sites frequently, but did not report their visit to the license holder.

The department realised that communication on the achievements needed improvement. As a consequence the transparency towards the license holder was increased.

- The benchmark showed that the department exceeded the time limits for issuing permits. Two actions were taken based on this insight from the performance comparison in the benchmark network. A person was appointed to support the department to bring down those figures (additional manpower) and a special license was introduced for less complicated permit requests that allows the applicant to receive a permit within one day.

The examples above show that dialogue and conjoint action can be observed in the municipality of Eindhoven, leading towards shared understandings (integration). Based on the benchmark results some new actions have been proposed and embedded (institutionalisation), such as process improvement and structural improved communication. So how was the knowledge transferred from the learning network to the municipal organisation?

Specific ad hoc task forces were established for issues derived from the benchmark report that required additional attention (i.e., the division of costs within the department). Respondents in the municipality of Eindhoven referred to the benefit of several external network environments, either being contact moments organised as part of the benchmark cycle, but also cross-organisational professional relationships and gatherings as organised by the Association of BWT. Respondent 21 observes the following: *"We get the most out of our contacts, because they tell you what they do. By watching where the differences are, you will end up learning"*. Respondents are of the opinion that the individual contacts with colleagues from other municipalities are most beneficial in learning from the benchmark. To have a conversation with each other allows for identifying learning processes and differences, posing in-depth questions and ultimately defining elements that can be smart to do as well. Respondent 24 is referring to a basic requirement for the identification process: *"You must dare to be vulnerable. But on the other hand you can learn a lot from it"*. The inter-organisational gatherings and the ad hoc (internal) task forces offer a platform for the feed forward of insights obtained in the benchmark network.

Emmen

In the municipality of Emmen, a few learning episodes of benchmarking could be detected from the interview excerpts that point towards a process of integration:

- The performance comparison showed that Emmen scored less on customer satisfaction with regard to information sharing. As a result communication has been improved.
- In comparison to other municipalities, the municipality of Emmen appeared to have a very high production level for issuing permits. At the same time it was observed that the quality of the permit review needed improvement. As a result, additional manpower was organised to lower the work pressure and improve the quality of the process.

It can be observed from the interview excerpts that a process through which shared understandings among individuals occur and coordinated actions through mutual adjustments are taking place (integration). The results of interpreting (communication of ideas to others) are dependent on the individuals involved and the environment within which the process occurs. Once the final benchmarking report arrives, the report is distributed to the heads of the sub-departments of the Building and Housing Department for further discussion. However, according to respondent 6, further discussion is not always taking place due to time constraints. Several other observations are made by the respondents why limited use is made of the benchmark information and why steps towards institutionalisation are limited.

> *"I can only speak for myself, my own team meeting. I put back a certain number of data, but you should be very careful with it. Especially within my team, team supervision, if you mention that we have a lot of staff members in comparison to other municipalities, then*

they do not understand, because they feel very busy. So, you should put it into a very good context" (respondent 4)

"At this moment, the benchmarking report comes ones a year. And based on that, you make adjustments. It seems like you are redirecting an oil tanker, while the organisation but also the society is so variable that you can no longer require government organisations to make adjustments on a yearly base. No, you need to make adjustments every day. So, in that sense it might be an out dated instrument" (respondent 6)

The few learning episodes of benchmarking detected reflect the ad hoc process of learning. The results of interpreting is dependent on individuals, however structured dialogue and routinised actions do not occur due to time constraints and perceived lack of benefit of the benchmarking results.

Groningen

In the municipality of Groningen limited learning episodes of benchmarking could be detected from the interview excerpts. According to the respondents the key benefit of the performance comparison is for purposes of strategic orientation and defensive use towards the mayor and municipal secretary in times of budget cuts. Instances of interpretation and institutionalisation could not be detected. It is claimed by the head of department that the indicators are not strong enough to carry out an organisational change and that suggestions for improvement come from their own analyses. It can be observed that the municipality of Groningen focuses more on exploitation than on exploration of the benchmark information. This is seen in the limited feed forward processes of learning that can be detected across the individual (in this case the head of department) towards the group and organisation levels. Respondent 19 observes: "*Actually very little is done with the benchmark. Much more can be done it is however the responsibility of the team leaders to follow up on the trends observed*". Table 5.1 gives a summary of claimed uses of benchmarking information within the municipal organisations mentioned during the interviews and described in the previous paragraphs.

Table 5.1 Claimed uses of benchmark information

Learning process	Claimed use	Case organisations			
		Breda	Eindhoven	Emmen	Groningen
Institutionalisation	System improvement	✓			
	Process improvement (routinised actions)	✓	✓		
Integration	Process improvement (shared understandings)	✓	✓	✓	
	Input for policy making	✓			
	Input for reorganisation	✓			
	Improved external communication	✓	✓	✓	
	Staff recruitment		✓	✓	
Interpretation	Strategic orientation				✓
	Defensive use		✓		✓

Source: the author

System improvement and process improvement (routinised actions) reflect institutionalisation. These types of improvements guarantee the embedding of individual and group learning within the organisation, including its symbolic and relational systems, structures, procedures and strategies. Examples were given in both Breda (improved permit registration system and turnaround times of the issuing of building permits) and Eindhoven (introduction of a special license that allows the applicant to receive a permit within one day). Process improvement (shared understandings), input for policy making and reorganisation, improved communication and staff recruitment reflect a process of integration. This can be observed when dialogue and conjoint action result in shared understandings. Examples were given in Breda, Eindhoven and Emmen. Comparison with others and defensive use can be seen as a personal explanation for one's self and for others about individual's insights (interpretation). In all the processes described patterns or possibilities are recognised from within an individual (intuition). The table below gives a summary of the processes of benchlearning.

Table 5.2 Processes of benchlearning that can be observed in the case studies

Process	Intuition	Interpretation	Integration	Institutionalisation
Breda	✓	✓	✓	✓
Eindhoven	✓	✓	✓	✓
Emmen	✓	✓	✓	
Groningen	✓	✓		

Source: the author

The empirical data suggest that in the municipalities where integration of the findings is taking place (Breda, Eindhoven and Emmen) much of the interpretation is done within the small circle of the management team and with peers in the benchmark network (and inter-organisational professional groups). In the case of Groningen, where a process of integration could not be detected, the benchmarking report is not shared within the department and not discussed within the management team. It can be concluded here that a combination of interpretative approaches at the group level with a mixture of hierarchical and network strategies, positively influences benchlearning.

5.3.4 Association of BWT

A feedback process can be observed from activities supported or initiated by the Association of BWT (or groups of municipalities) feeding directly or indirectly into the municipal organisation. Although the feedback process from the Association of BWT towards the municipal organisation has not been the focus of the empirical data collection, some observations can be shared in this section.

The municipalities that participated both in the platform and the benchmark since its establishment in 1999 included (among 10 other municipalities): Breda, Eindhoven and Emmen. The municipalities of Breda and Eindhoven have been actively involved in the development of the benchmark network and are active members of the quality improvement initiatives in the building and housing sector. Lessons that can be detected are grouped around the three perspectives of the benchmark questionnaire: the client perspective (quality of service delivery); the management perspective (effectiveness) and the organisation perspective (efficiency)[1].

Client dimension:

- A group of municipalities (under the umbrella of the Association of BWT) started a joint customer satisfaction survey, a dimension they were missing in the benchmark. This later turned into a professional instrument (managed by a private company). The benchmark data and satisfaction surveys offered new insights about the importance that customers attach to consultation and the municipality being accessible. Based on these insights several municipalities started to work based on appointments and built in forms of consultation in their process.
- A project, initiated by the Association of BWT and financed by a group of large municipalities, to build a central server for building permits, something individual municipalities could not have done by themselves. This initiative (supported by the Association of BWT) has been adopted by central government and the online request for building permits has been formalised.
- Other activities that have been supported and/or initiated by the Association of BWT on request of the participating municipalities in the benchmark are: the development of a knowledge base and forum for building supervisors (BWTinfo.nl) and the launch of a professional magazine.

[1] As discussed with Dr Marcel Hoogwout (Zenc) in an e-mail correspondence on the 20th of May 2012.

Management dimension:

- From the first benchmark rounds it became clear that the understanding and comparability of quality standards for building permits was not uniform. As a result, the national key protocol for the collective quality norms for building permits (LTP – previously called CKB) was launched with the assistance of the Association of BWT. The central government raised the key protocol to a minimum quality standard for individual municipalities and other regulators.
- A similar initiative was launched with the assistance of the Association of BWT (initiated from the municipalities participating in the benchmark) to monitor the quality of the supervision on the construction site by municipalities. The monitoring protocol serves for normalizing and comparing the quality by municipalities. The monitoring protocol will be combined with monitoring protocols from other supervisors (such as the environment and the fire brigade) to increase the efficiency and effectiveness.
- The discussions taking place in the learning group as part of the benchmark cycle also resulted in fundamental discussions about the functioning of the building regulations. The Association of BWT developed itself as the interlocutor between the lawmakers and the municipalities.

Organisation dimension:

- The benchmark made clear that the interpretation of definitions was not standardised among municipal representatives. Already after the first round of the benchmark, a group of municipalities started to standardise the costing of fees. In addition to this, a call was made by the Association of BWT, to standardise the individual municipal administrations and account schedules, in order to facilitate the comparison and standardised interpretation.
- The need to standardise fee calculations became apparent among builders and developers. To date, several working groups are active dealing with this subject with the support from the Association of BWT.

It can be observed that (in addition to the feedback loop from the individual to the municipal organisation) a feedback loop is taking place from the Association of BWT feeding directly and/or indirectly into the municipal organisation. Through the group process of dialogue and conversation the group members enabled a deeper meaning of the issues at stake. A map slowly emerged with finer and finer levels of detail. Several of these have resulted in coordinated actions benefiting all municipalities. It can be observed that coordinated actions are facilitated by the Association of BWT, either initiated through the benchmark network or a separate group of municipalities. What has already been learned, feeds back from the Association (directly or indirectly) to the municipal organisation to group and individual levels, affecting the work of the Building and Housing Department. Several examples refer to initiatives taken by groups of municipalities and sometimes there is not a direct link from the Association to the municipal organisation, as the central government is involved as well.

5.4 Learning barriers

In this section, an attempt is made to explain the differences in processes of benchlearning that can be observed in the case studies (see table 5.2). What are the learning barriers that complicate or impede the learning to move from one analysis level to the other? Schilling and Kluge (2009) categorise and discuss different learning barriers in the form of factors complicating or impeding organisational learning based on the 4i model developed by Crossan *et al*. (1999) and further developed by Lawrence *et al*. (2005). This section uses the learning barriers framework developed by Schilling and Kluge (2009) to classify the empirical data from the four case studies. Learning barriers are defined as "those factors either preventing organisational learning or impeding its practicability" (Schilling and Kluge, 2009, p. 337). The authors define different forms of learning barriers: personal-action barriers (characterised by individual thinking, attitudes and behaviour); structural-organisational barriers (organisational strategy, technology, culture and formal regulations); and societal-environmental barriers (clients, suppliers, competitors, socio-political environment and technology). A fourth level of learning (benchmark network) is added to the learning barriers' framework based on the analysis as presented in the previous section. The related learning barrier can be formulated as the inter-organisational barrier (characterised by the network structure and knowledge transfer). This section attempts to shed more light on what actually makes the moving between the learning levels problematic. Which factors hinder the moving from one step to another? Insight in the learning barriers may assist practitioners in the field of benchmarking to develop strategies and tactics to counter such factors.

Table 5.3 Barriers to benchlearning

Form	**Municipality**	**Barriers**
Inter-organisational	Breda	Different interpretation of the indicators (unreliable data)
	Eindhoven	Incomplete data (giving a wrong impression) Frequent introduction of new members to the learning network Traditional versus progressive municipalities in the benchmark network The difference in municipal size makes comparison difficult
	Emmen	Rotation of municipalities participating in the learning network
	Groningen	Knowing the context is important for making a good analysis The benchmark gives only basic information The benchmark results do not give direction for improvement Negative perception of the benefit of benchmarking
Action-personal	Breda	Lack of political and social skills on part of the contact person Lack of dissemination from the side of the manager Occupied by issues of the day
	Eindhoven	Hard to push the right button due to the level of complexity
	Emmen	Time constraint Personal attitudes and behaviour Lack of top management support
	Groningen	Lack of motivation to excel
Structural-organisational	Breda	Lack of resources to initiate change process Lack of clear responsibility concerning the implementation Inconsistency between interests of different internal stakeholders Benchmarking is not included in the planning and control cycle Lack of political interest in the benchmark Benchmark report offers out dated information
	Eindhoven	-
	Emmen	Delinked from the primary process Inconsistent organisational systems Separation between the operational and the policy departments

		Lack of reliability of the internal performance information system Inward focus Lack of clear goals Lack of clear responsibility concerning data storage
	Groningen	Lack of staff resources High workload of the management team and/or staff members
Societal-environmental	Breda	Assertive community (no time for learning)
	Eindhoven	-
	Emmen	Unclear criteria for success formulated in the sector
	Groningen	Decrease in income from permits (due to a decrease in building activities)

Source: the author

A long list of barriers could be detected within the case organisations that complicate or impede the benchlearning to move from the group towards the organisation level. Some of the reasons mentioned are: lack of clear responsibility, inconsistency between interests of different stakeholders, lack of clear goals, delinked from the primary process etc.

As could be observed in section 5.3, learning processes in Emmen and Groningen have developed less than in the other two case municipalities. The barriers to benchlearning as presented above give some insight into possible reasons for this. In the case of Groningen it can be observed that the barriers refer to issues of motivation, a perceived lack of relevancy of the outcomes (only basic information, do not give direction for improvement) and work pressure. In the case of Emmen it can be observed that the majority of the barriers refer to structural-organisational issues. At the same time it can be observed that the municipality of Breda and Eindhoven as well refer to a long list of barriers on most analysis levels. The barriers as presented to benchlearning are hence not necessarily a decisive reason for limited learning across all analysis levels. In the next section supporting factors are listed which shed more light on the success of both Breda and Eindhoven in moving the learning from one analysis level to the other.

5.5 Supporting factors

Schilling and Kluge (2009, p. 356) observe that "*pulling down barriers is a necessary, but not all-sufficient, condition for organisational learning, as removing the factors that prevent learning is only the first step in supporting organisational learning*". The authors suggest an analysis of supporting factors for organisational learning. The table below gives insight into the *supporting* factors for benchlearning as mentioned by the interviewees.

Table 5.4 Supporting factors to benchlearning

Form	**Municipality**	**Supporting factors**
Inter-organisational		-
Action-personal	Breda	Communication skills of the individual members of the management team Pro-activeness of the individual
	Eindhoven	Intrinsic motivation to improve
	Emmen	-
	Groningen	-
Structural-organisational	Breda	Departmental communication and cooperation Team work Specialists in each team
	Eindhoven	Shorter hierarchical lines and organic teams
	Emmen	Launching of an improvement team to work on a specific issue
	Groningen	-
Societal-environmental	Breda	Show commitment, listen to people and search for dialogue with citizens for a better understanding of the issues at stake Participation in external networks
	Eindhoven	Participation in external networks
	Emmen	-
	Groningen	-

Source: the author

Three observations can be made from the listing above and the additional case analysis done. A first observation is that according to the respondents, supporting factors for benchlearning should be predominantly looked for at the individual, group and organisational level. The inter-organisational level serves as an initial step for comparison with other municipalities and serves as an initial platform for interpretation. The respondents observe that the aggregated and individual benchmarking report gives sufficient information to act upon and that the responsibility for follow-up lies predominantly with the organisation itself.

The second observation is that the empirical data showed very limited reference to supporting factors towards benchlearning from interviewees in the municipalities of Emmen and Groningen. Interviewees from Breda and Eindhoven referred to supporting factors at the action-personal, structural-organisational and societal-environmental level. The supporting factors suggest that in order to facilitate benchlearning intrinsic motivation and communication skills matter, supported by a high level of cooperation (team work) and a flat organisational structure. The participation in external networks and a continuous search for dialogue appear to have a positive influence on benchlearning.

A third observation is that most supporting factors refer to the interaction between individuals. This suggests that municipal benchmarking is likely to deliver improved organisational performance when it is based on interactions between individuals.

5.6 Conclusion

The organisational learning theory provided a useful lens to understand why some municipalities are better able to learn, while others only embrace some of the useful benchmark information. The adjusted 4i framework of organisational learning (Crossan *et al.*, 1999; Dutta and Crossan, 2003) and the learning barriers' framework (Schilling and Kluge, 2009) allowed for structuring the empirical data.

The *first* research question of this chapter looked into the analysis levels and related learning processes that can be observed in municipal benchmarking. Although the 4i framework provides a compelling basis for research on organisational learning, it neglects the inter-organisational level present in municipal benchmarking. Consequently, a wider conception of learning is suggested for benchmarking, where learning processes interact with knowledge exploitation and exploration dynamics throughout four analysis levels (i.e. inter-organisational, individual, group and organisational levels) in two parallel communities (benchmark network and municipality).

A hybrid approach could be observed in the more successful municipal organisations (Breda and Eindhoven) combining interpretative approaches at the group level with a mixture of hierarchical and network strategies to move learning from one analysis level to the other. The sharing of understanding with the members of the group is reached through interactions between individuals steered at the initiative of the head of department (who also attends the benchmark network). Action is taken by the department heads in Breda and Eindhoven, leading to improvements in the sector, without necessarily sharing their understanding with the operational and the tactical level within the department. It could be observed that in the case authority where no examples of integrating the benchmarking information could be observed (Groningen) interpretative processes at the group level within the municipal organisation was absent. It appeared difficult to shift from individual learning to learning among individuals or groups in this municipality. This suggests that moving from interpreting to integrating (feed forward) can be seen as a problematic interaction that can be supported by a hybrid approach combining interpretative approaches at group level with a mixture of hierarchical and network strategies.

It can be concluded that participation in external networks impact upon the learning process and the organisational ability to mobilise knowledge. This observation suggests that actions taken within the municipalities are not clearly separated from actions taken within the network.

It can be observed that coordinated actions are facilitated by the Association of BWT, either initiated through the benchmark network or through a separate group of municipalities. What has already been learned, feeds back from the Association (directly or indirectly) to the municipal organisation and to the group and individual levels, affecting the work of the Building and Housing Department. Several examples refer to initiatives taken by groups of municipalities and sometimes there is not a direct link from the Association to the municipal organisation, as the central government is involved as well. Overall, it can be observed that the situation in practice is more complicated than initially suggested. The inter-organisational level within the scope of this research study should not refer to the benchmark network, but to a broad arena of inter-organisational groups (groups of municipalities, professional groups etc.).

It can be acknowledged that the routines developed within the benchmark network (goals are clearly formulated, interests of stakeholders are similar, responsibilities are clear within the network, an annual planning for the benchmark cycle and the meetings of the learning network is available) and the advocacy role taken up by the Association (towards national government) is beneficial for the institutionalisation of the outcomes. A similar institutionalised system to exploit the benchmark information is absent in all four municipalities. The exploitation of the results is based on ad hoc interventions by the head of department and is not aligned to any system, structure or formal mechanism.

Analysis of the *second* and *third* research question offered insight into the factors that hinder or support the moving from one step to another within the learning process of benchmarking. It can be observed from the empirical data presented that the barriers to benchlearning are not necessarily a decisive cause of limited learning across all analysis levels. From the analysis it can be concluded, however, that supporting factors have a strong impact on moving the learning from one analysis level to the other. In order to facilitate benchlearning intrinsic motivation and communication skills matter, supported by a high level of cooperation (team work), a flat organisational structure and interactions between individuals. The participation in external networks and a continuous search for dialogue appear to have a positive influence on benchlearning. Hence, stimulating supporting factors can be seen as a more important strategy to stimulate benchlearning than pulling down barriers.

6. Power and politics of benchlearning

"Power and status differences are pervasive in organisations and have important and far-reaching implications for perceptions, motivations, and behaviour. Yet past research on learning in organisations and groups has tended to overlook the implications of power and status differences, building instead from a set of rational system assumptions that may be inappropriate for settings where power and status differences exist".

(Bunderson and Reagans, 2010, p.11)

6.1 Introduction

This chapter gives insight in the forms of strategic stances and power that can be seen in the cases and how these different forms affect the learning outcome of benchmarking. The learning outcome can be classified as single-loop learning (target narrow process improvement) or double-loop learning (a broad understanding of policy choices and effectiveness). The question of whether there is a relationship between the strategic stances adopted and the forms of power identified in the four cases will be examined.

In understanding and evaluating the impact of municipal benchmarking, it is important to understand where and how impact is created. Much of the 'why' of change is derived from an analysis of the *context*. This means that it is important to interpret and model the context within which impacts occur (Pettigrew, 1990). Coopey (1995, p. 2004) observes that since knowledge is such an important resource in a learning organisation, much political activity is expected to be associated with how it is acquired and stored. In addition to this, Brignall and Ballantine (2004, p. 231) observe that an institutional approach explicitly "recognises the importance of relative bargaining power in determining whose interests will predominate in an organisation and the consequent effects on what aspects of performance are measured, reported and acted upon".

To have a better insight in the forms of strategic stances and power that can be seen in the cases and how these different forms affect the learning outcome of benchmarking, the following research questions are explored in this chapter giving insight in the *inner* context (the structure, culture, and political context within the organisation through which ideas for change have to proceed) of the case municipalities:

- Which forms of strategic stances can be seen in the four cases?
- Which forms of power can be seen in the four cases?
- How do different forms of strategic stances and power affect the learning outcome of benchmarking in the building and housing sector in the Netherlands?

This chapter relies on two typological classifications to explore the research questions. Firstly, the typological classification of Miles *et al*. (1978) is used to analyse the forms of strategic stances. Their classification suggests that to solve their problems, organisations employ four strategic types: defenders, analysers, prospectors and reactors. Secondly, this chapter relies on the conceptualisation and typology of power as formulated by Mintzberg and Waters (1985), Lawrence *et al*. (2005, 2012) and Lawrence and Robinson (2007). Lawrence *et al*. (2005) integrated power into the 4i framework of organisational learning as presented in the previous chapter. A slightly modified version of their framework is well suited for the analysis of the empirical data of this chapter because it relates forms of power in organisations to specific learning processes.

This chapter examines the research questions based on the 24 semi-structured interviews conducted between November 2009 and September 2010 in four Dutch municipalities. To analyse the data two approaches were employed. To gain familiarity with the data and to accelerate cross-case analysis, a pattern of thought in each individual case was sought for by doing a within-case analysis. To compare the parameters of each case and to find common thoughts among the cases a cross-case analysis was done. Interview quotations referring to the strategic stance of and the forms of power in the municipalities were identified to determine and analyse the strategic stance of each case study and the forms of power present at the time of the interview.

Section 6.2 gives a brief introduction into the forms of strategic stances and how these relate to organisational learning. The typological classification of Miles *et al*. (1978) is presented in this section. The subsequent section (6.3) gives insight into the forms of power and how these relate to organisational learning. The empirical data related to the research questions are presented in section 6.4 followed by a cross-case analysis in section 6.5. Some concluding observations are presented in section 6.6.

6.2 Strategic stances and organisational learning

In this section insight is given into the relationship between an organisation's strategic stance and its social and psychological processes of organisational learning. Dutta and Crossan (2003, p. 17) note that the organisational learning literature has only recently recognised the importance of power and political processes and their impact on learning. The authors continue to observe that in the change literature, however, discussion of these organisational forces had been incorporated several years back. Pettigrew (1987, p. 666) observed that: "*… the development of strategic change in the firm takes on the character of a political learning process, a long-term conditioning and influence process designed to establish the dominating legitimacy of a different pattern of relation between strategic content, context, and process*."

Andrews *et al*. (2006) suggest that measures of strategy content must be included in valid theoretical and empirical models of organisational performance in the public sector. In the definition used by the authors, strategy content comprises two dimensions: strategic stance (the extent to which an organisation is a prospector, defender, reactor or analyser – see Miles *et al*., 1978) and strategic actions (the relative emphasis on changes in markets, services, revenues, external relationships, and internal characteristics).

Back in 1978, Miles et al developed a general model of the adaptive change process which they called the "adaptive cycle". The cycle is based on the strategic-choice perspective. Miles *et al*. (1978, p. 548) describe the essence of the strategic-choice perspective: "*organisational behaviour is only partially preordained by environmental conditions and that the choices which top managers make are the critical determinants of organisational structure and process*". To solve their problems, organisations employ four categories of strategic types: defenders, analysers, prospectors and reactors. Although Miles et al indicate that any typology is unlikely to encompass every form of organisational behaviour, every organisation fits predominantly into one of the four categories. The central contention of Miles *et al*. (1978) is that prospectors, defenders, and analysers perform better than reactors. Gupta (2011, pp. 512-513) - as well based on other authors - describes the four typologies as follows:

> *Defenders* are internally oriented organisations. They stress efficiency, and are tightly organised firms focused on maintaining a niche with a limited range of products or services. These firms devote primary attention to improving the efficiency of existing operations. They develop a core technology that is highly efficient and use an organisation structure with centralised control.
>
> *Prospectors* are the exact opposite of defenders. The prospectors have an external focus. They are continuously involved in monitoring the external environment as they intend to respond quickly to early signs of opportunities and to exploit the benefits of being a pioneer in a new product/market area. The structure of these firms is characterised by a low degree of formalisation and routine, decentralisation and lateral as well as vertical communication, emphasising aspects such as innovation and flexibility.
>
> *Analysers* blend the characteristics of prospectors and defenders. The analyser maintains a moderate level of business risk by waiting to see the experience of others before entering a market. They put emphasis on longer-term planning and much thought about decisions prior to action in most instances. The analyser partitions its technology so that it can serve its stable domains with efficient technologies and its dynamic domains with flexible and effective technologies. They include flexibility as well as stability, adopting structures that can accommodate both stable and changing domains.
>
> *Reactor* organisations do not present any consistent pattern of response behaviour to environmental conditions. Their actions are mostly reactions to outside forces,

such as the economy, competitors, or market pressures. They lack a consistent strategy-structure relationship. They are not planners, but reactive and thinkers by necessity.

The multiple-informant survey of 119 English municipalities implemented by Andrews *et al*. (2006) shows that organisational performance is positively associated with a prospector stance and negatively with a reactor stance. Meier *et al*. (2007) find that the effectiveness of a strategic stance is related to the specific goal aimed for. The authors find that the defender strategy is the most effective for the primary mission of the organisation and that the prospector and reactor strategies work best in regard to the goals of the more politically powerful elements of the organisation's environment.

Mintzberg and Waters (1985) conclude in their paper on deliberate and emergent strategies, that strategy formation walks on two feet: one deliberate and the other emergent. Deliberate and emergent strategies may be conceived as two ends of a continuum along which real-world strategies lie. Whereas the more deliberate strategies tend to emphasise central direction and hierarchy, the more emergent ones open the way for collective action and convergent behaviour. In their view, the fundamental difference between a deliberate and an emergent strategy is that whereas the former focuses on direction and control, getting desired things done, the latter opens up this notion of 'strategic learning'. The authors suggest that adding the concept of emergent strategy opens the process of strategy making up to the notion of learning. Emergent strategy implies learning what works, taking one action at a time in search for that viable pattern or consistency.

Relating the view of Mintzberg and Waters (1985) to the four strategic stances as suggested by Miles *et al*. (1978), it can be observed that defenders define strategy as deliberate, devoting primary attention to improving the efficiency of existing operations using an organisation structure with centralised control. Prospectors on the other hand define strategy as emergent. They are continuously involved in monitoring the external environment as these intend to respond quickly to early signs of opportunities and benefits. Analysers blend the characteristics of both the prospector and defender orientations. They include flexibility as well as stability, adopting structures that can accommodate both stable and changing domains. As such, analysers blend the characteristics of both deliberate and emergent strategies. Reactors do not present any consistent pattern of either deliberate or emergent strategies, lacking a consistent strategy-structure relationship.

Bringing strategic stances into research on municipal benchmarking should provide a more effective foundation for understanding the learning outcome from participation in the benchmark.

6.3 Power and organisational learning

Power is a complex concept with a multitude of definitions and connotations in the literature. In the context of the developing power debates of the 1950s to the 1970s, it has frequently been stated that power is best characterised as an 'essentially contested' concept (Haugaard, 2010, p. 419). Lawrence *et al*. (2012, p. 102) observe that organisational research in the 1970s and 1980s made power a central element in the study of change, demonstrating the importance of the political strategies of organisational actors. The authors refer to a distinction made by Clegg *et al*. (2006):

> "*Distinctions among conceptions of power include that between "power to" and "power over". The former points to an understanding of power as facilitative, allowing one to do something one otherwise would not be able to do – the power to vote, the power to access information, etc. The latter highlights power as advantage – power as a restrictive mechanism through which one actor controls the potential actions of another – the power of a boss over a subordinate, the power of a parent over a child*".

Haugaard (2010) proposes a plural view of power, consisting of a cluster of concepts, each of which qualifies as 'power'. Haugaard (2010, p. 420) argues: (i) that power is a family resemblance concept, which entails that there is no single 'best' definition of power and

(ii) that these family members can legitimately change their meaning depending upon which *language games* are being played.

The inclusion of power and politics helps to explain why some organisations are more successful at learning. Lawrence *et al*. (2005, p. 182) argue that "the distinction between episodic and systemic forms of power can help us understand how organisational politics affect the movement of ideas from individuals through groups to the organisation, and from the organisation back to individuals". The authors connect episodic power to the movement of ideas "upward" in organisations, from individuals to groups or the organisation as a whole; they associate systemic forms of power with the movement of ideas and practices "downward" from the organisation as a whole to specific groups and individuals. Holmqvist (2004, p. 72) argues that in order for an organisation to shift to either exploration or exploitation, a challenge to the views and assumptions held by dominant organisational actors is required. These challenging individuals (or challenging group) need to be able to identify and articulate to the dominant organisational actors (i.e. strategic level) a problem that needs a new approach in order to be solved.

Lawrence *et al*. (2005, p. 189) suggest that organisations which lack balance between episodic and systemic power will also fail to manage the exploitation (using what has already been learned) / exploration (assimilating new learning) tension adequately: organisations where the expression of episodic power is inadequately nurtured will suffer from a lack of exploration, whereas organisations which are characterised by relatively weak systemic power will perform poorly in terms of exploitation. Coopey and Burgoyne (2000) argue that organisational learning is facilitated by a free and open form of politics. The authors suggest that an open form of politics produces an organisational environment that stimulates individuals and groups to take up the challenge of experimenting with new alternatives (unlearning). Strategies for unlearning existing knowledge and for reframing what is already known are constrained by entrenched power structures and the associated patterns of dependency. Coopey and Burgoyne (2000, p. 882) argue that: "unless political action enables these structures to be challenged, higher-level learning will be inhibited".

The empirical results from a nationwide Norwegian benchmarking project offer indications that learning from benchmarking is subject to politics (Askim *et al*., 2007). Factors such as network and administrative characteristics and management and political participation are found to influence learning outcomes.

Lawrence *et al*. (2005) have integrated power into the 4i framework of organisational learning that was introduced in section 1.3.2 and further explored in chapter 5. The authors argue that power and politics provide the social energy that transforms the insights of individuals and groups into the institutions of an organisation. The political will and skill of those attempting to make a transformation of new ideas into coherent collective action is a key issue in understanding which ideas will be integrated into the activities of groups and which will become institutionalised in organisations. March (1991) points out that organisations need both to exploit existing ideas and opportunities, and to explore new ones, to be successful in changing environments. While exploration is associated with the notion of learning experiences made by individuals and groups, exploitation occurs by including the aspect of transferring the experiences into organisational routines, structures and processes.

Lawrence *et al*. (2005) propose that different forms of power in organisations are connected to specific learning processes – intuition is linked with discipline, interpretation with influence, integration with force, and institutionalisation with domination – and that an examination of these different forms of power provides a basis for understanding why some insights become institutionalised while others do not. The episodic forms of power can be described as *influence* and *force* and the systemic forms as *domination* and *discipline*.

Discipline (linked with intuition) supports and shapes the intuitions of organisational members by providing them with an on-going stream of experience and affecting the ways in which they perceive that experience by shaping their identities. It involves practices such as recruitment, socialisation, compensation, training and team-based work. *Influence* (linked with interpretation) is affecting the costs and benefits that organisational members' associate with specific interpretations of a new idea. The use of influence in affecting interpretation can involve a wide range of tactics, including moral suasion, negotiation, persuasion, ingratiation and exchange. The process of *force* (linked with integration) is characterised by creating circumstances that restrict the options available to organisational members using formal authority to implement the new idea. This might involve restricting the consideration of alternative practices, restricting issues for discussion on formal and informal agendas, and removing/transferring opponents of the innovation. To overcome particular resistance to institutionalising changes, *domination* (linked with institutionalisation) is regarded as a particularly effective strategy. The importance of domination as a political basis for institutionalisation is tied to the manner in which it addresses potential resistance to change. Systems of domination do not rely on the potentially unreliable or unpredictable choices of individuals. Instead, these systems manage organisational stakeholders "in place", thereby affecting their behaviours without necessarily shaping their preferences, attitudes or beliefs. Although the term domination might have negative connotations and potentially suggest a coercive environment, it is not meant to imply that it is simply repressive. Systems of domination can be enabling, productive and enjoyable for the members they affect (Schilling and Kluge, 2009, pp. 184-188).

The four forms of power occur across the learning levels. Examples for each form of power are given in table 6.1. Connecting episodic power to feed-forward learning processes and systemic power to feedback learning processes provides a foundation for understanding how organisational politics influence benchlearning. Although the definitions are presented in a linear fashion in the table below, it is critical to appreciate the iterative nature of the forms of power. Discipline occurs at the individual level and domination at the organisational level; however, influence bridges the individual and group levels, while force links the group and organisational levels. See the table below for an overview of the relationships Lawrence *et al*. (2005) are proposing.

Table 6.1 The politics of organisational learning

Process	**Intuition**	**Interpretation**	**Integration**	**Institutionalisation**
Definition	The "recognition of the pattern and/or possibilities inherent in a personal stream of experience"	The process of "explaining, through words and/or actions, … an insight to one's self and to others"	The process of "developing shared understanding among individuals and of taking coordinated action"	The process of "embedding learning that has occurred by individuals and groups into the organisation"
Key requirements	Developing organisational members with experience and identities that facilitate expert-level pattern recognition	Managing ambiguity and uncertainty in the adoption of language and the construction of cognitive maps	Translating new ideas consistently across members in order to achieve collective action	Overcoming the resistance to change of organisational members
Associated form of power	**Discipline**	**Influence**	**Force**	**Domination**
How it works	Affecting costs / benefits of behaviours	Affecting costs / benefits of behaviours	Restricting available behaviours	Restricting available behaviours
Examples	Socialisation Training Team-based work	Moral suasion Negotiation Ingratiation	Agenda setting Limiting decision alternatives Removing opponents	Changing material technologies Changing information systems Changing physical layout

Source: adapted and modified from Lawrence *et al.*, 2005, p. 185

The forms of strategic stances and power as presented in the previous sections are taken as the entrance point for analysing the case studies in section 6.4. Pettigrew (1992) observes that context and action are always interwoven: "*context is not just a stimulus environment, but a nested arrangement of structures and processes where the subjective interpretations of actors perceiving, learning and remembering help shape process*" (p. 10). The work of Pettigrew (1987, 1992 and 1997) is guided by the precept that the exploration of strategic choices and changes has to be embedded in an analysis of the inner and outer context of the organisation. This means that it is important to interpret and model the context within which learning outcomes from benchmarking are shaped and steered. This approach is supported by Brignall and Modell (2000, p. 282) who suggest that "*it would appear appropriate to shift the attention to the power and pressures exerted by different groups of stakeholders and how these affect the use of performance information in organisations*".

6.4 Case descriptions

In this section insight is given in the forms of strategic stances and power that can be seen in the cases and how these different forms affect the learning outcome of benchmarking. The section starts with a description per case to address the form of strategic stance (defender, prospector, analyser and reactor), followed by a description of forms of power that can be seen in each municipality (discipline, influence, force and domination). In the subsequent section (6.5) a cross-case comparison is done to analyse how these different forms affect the learning outcome of benchmarking.

6.4.1 Breda

From a strategic stance perspective, the municipality of Breda can be seen as an *analyser*. Breda shows a combination of prospector and defender characteristics. To accommodate both stable (defender characteristic) and dynamic (prospector characteristic) areas of operation, the municipality of Breda opted for a matrix organisation. The matrix structure that can be observed combines both functional divisions and product groups. If work pressure is unequally divided, a flexible work attitude can be observed where colleagues help each other out. The administrative structure in Breda can be characterised by a proper differentiation of the organisation's structure and processes to achieve a balance between the stable and dynamic areas of operation. Examples can be found where best practices by other municipalities are adopted (e.g., capturing the statutory building regulations into policies), a characteristic of the analyser stance, imitating demonstrable successful approaches. They are continuously involved in monitoring the external environment and explore and discuss new developments in the field characterised by a moderately centralised control.

Which forms of power can be seen in the Building and Housing Department of the municipality of Breda possibly influencing the learning outcomes of benchmarking? *Discipline* can be observed (systemic mode) in terms of socialisation, training opportunities and team-based work. Several actions were taken in the municipality of Breda to shaping the actual formation of staff members' perceptions. In Breda, appreciative innovation can be seen as a new approach to communication and team work. Respondents referred to the embedded learning that has occurred within the Building and Housing Department as a result of the introduction of appreciative innovation. The new approach to communication and collective action gives a strong foundation for benchlearning within the department. The respondents indicated that the benchmarking report can be used to *influence* the council and the alderman. The comparison with other municipalities and the results over time as presented in the benchmarking report are seen as useful tools to affect interpretation in situations in which improved action, decrease in costs and increase in benefits of the department is required from the council and the alderman.

6.4.2 Eindhoven

From a strategic stance perspective, the municipality of Eindhoven fits predominantly in the *prospector* typology. The municipality went through several reorganisation processes over the past years. Miles *et al*. (1978, p. 553) observe that the prospector requires a good deal of flexibility in its technology and administrative system. An example is the decision to disconnect a program focused on improved service delivery (including five program lines) from the introduction of the WABO, in order to be flexible and to support continuous development. Another example of this is the quick response of the municipality to reduce the high number of permits that were issued after the deadline. Eindhoven invested in an individual who could act quickly and flexibly on all aspects of the permit cycle if required. A last example is the introduction of the "flash permit" as one of the first municipalities in the Netherlands. The flash permit allows the applicant to receive a permit within one day. Eindhoven can be seen as a creator of change and as a consequence deals with more uncertainty and reorganisations than municipalities with other stances.

Which forms of power can be seen in the Building and Housing Department of the municipality of Eindhoven possibly influencing the learning outcomes of benchmarking? In Eindhoven it could be observed that the dominance of the municipal executive team prevented the departmental management from being transparent about the results of the more recent benchmarks. Interviewees indicated that sharing the benchmarking results with the alderman has already resulted twice in a lot of trouble (*force*). The alderman could not accept the results in the reports. Due to this the feed forward of the benchmarking report from the group to the organisational level has been obstructed. The motivation of the reorganisation in 2009 can be found in the intended introduction of multiple permits (WABO) in 2011. This resulted in the merger of several departments (permit, supervision and enforcement) which improved the communication between departments. The sector head has given it a more organic course which first started by *force*, which led to a lot of resistance. A more systemic approach was used to overcome this and changes were made in the physical layout of the department (*domination*). Teams were brought together on the same office floor, physically reducing distance. At that moment steering and team building improved.

6.4.3 Emmen

From a strategic stance perspective, the municipality of Emmen lacks a consistent strategy-structure relationship and can be typified as a *reactor* in the typological classification of Miles et al at the time of the interviews. Emmen did not show a consistent pattern of response behaviour to environmental conditions. After the start of the reorganisation in 2008 (introduction of flexible working) management had articulated a direction for the organisation which involved operating in a cost-efficient way. It appeared that the Building and Housing Department did not re-design the structure and processes to fit the chosen flexible working strategy. One of the departmental managers pointed to the consequences of the reactive attitude of the municipality to solve issues:

"The hiring of externals is about 80% of the organisation's personnel costs, if not more…" (respondent 3). Interviewees mention the absence of a strategic view and the reactive attitude to issues of the day at the management level. According to Miles *et al*. (1978), the reactor is a "residual" strategy, arising when one of the other three strategies (defender, analyser or prospector) is improperly pursued. Two other reasons have been identified by Miles *et al*. (1978, pp. 557-558) why organisations become reactors: top management may not have clearly articulated the organisation's strategy; and a tendency for management to maintain the organisation's current strategy-structure relationship despite overwhelming changes in environmental conditions. In addition to the inappropriate link of the structure-process features, interviewees believe that the management of the Building and Housing Department did not articulate clearly the organisation's strategy in reaction to the cost-efficient direction articulated by the municipal management.

Which forms of power can be seen in the Building and Housing Department of the municipality of Emmen possibly influencing the learning outcomes of benchmarking? The municipality is primarily overcoming the resistance to change of organisational members by *domination*. Behaviours are restricted by introducing changes in the physical layout of the organisation. Episodic forms of power could not be detected, resulting in the failure to manage exploration dynamics. The municipality went through an organisational change process (some respondents talk about a reorganisation process) in 2008. The Building and Housing Department had not designed the structure and processes to fit the chosen flex working strategy. Due to this a severe backlog was created. A head of department confirms that the introduction of flex working within the Building and Housing Department resulted in a lower production and as a consequence resulted in delays in the license application. The absence of information systems supporting the introduction of the organisational change process in general and flex working in particular is perceived as a serious bottleneck by the interviewees. At the same time, the interviewees did not show the willingness or capability to influence organisational decision making. According to the literature, the results of interpreting (communication of ideas to others) are dependent on the individuals involved and the environment within which the process occurs. Lawrence *et al*. (2005, p. 182) argue in this respect that "the ideas that are successfully transformed into such interpretations will depend significantly on the episodic power of the ideas' sponsors – their ability to influence the thoughts, feelings, and behaviours of those around them". The municipality appears to have a group of hard working employees, however political behaviour is missing. The observations and ideas as generated in the benchmarking report lack active, interested members who are willing or capable to engage in political behaviour that pushes ideas forward and ensures their interpretation, integration and institutionalisation.

6.4.4 Groningen

From a strategic stance perspective, the municipality of Groningen can be typified as a *defender*. The Building and Housing Department is "*the richest, largest and most powerful department in the municipality*" (respondent 19). Due to this the department has a strong hold on the city. The department produces a limited set of products directed at a narrow segment of the total potential market (for example student housing). According to Miles *et al*. (1978, p. 551), defenders "tend to ignore developments and trends outside of their domains, choosing instead to grow through market penetration and perhaps some limited product development". It can be observed that Groningen maintains strict control in order to ensure efficiency. Little scanning is done of the environment and the management team is dominated by cost-control specialists. The process of planning receives a lot of attention in Groningen toward cost and other efficiency issues. An individual has been appointed to describe and monitor all process steps of the permit procedure. This great deal of investment in delivering services as efficiently as possible resulted in the fact that the department is "*already more than three years 100% covering its costs*"(respondent 19).

Which forms of power can be seen in the Building and Housing Department of the municipality of Groningen possibly influencing the learning outcomes of benchmarking? The forms of power observed in Groningen are primarily episodic (*influence* and *force*). Interpretation of the benchmarking report is done by the head of the Building and Housing Department in isolation and not shared with others. From the interview excerpts is could be observed that the head of the Building and Housing Department influences which issues derived from the benchmarking report are put forward for discussion on formal and informal agendas (restricting issues). The interviewees claim that changes come from their own analyses of daily happenings. The use of force as a form of power in Groningen is tied closely to formal organisational hierarchies. The benchmarking report is not being used as a steering instrument or as an input to annual plans or other departmental items. The influence and force that can be observed in Groningen is resulting in imbalances of power and status within the municipality.

6.5 Cross-case analysis

This section offers a cross-case analysis of how different strategic stances and power affect benchlearning. The table below (6.2) summarises the episodes of power that could be detected within the Building and Housing Departments of the four case studies as presented in the previous section. The results as presented in chapter 5 have been integrated in the table below for the sake of analysis.

Table 6.2 The politics of benchlearning in the four case studies

Process	**Intuition**	**Interpretation**	**Integration**	**Institutionalisation**
Breda	✓	✓	✓	✓
Eindhoven	✓	✓	✓	✓
Emmen	✓	✓	✓	
Groningen	✓	✓		
Forms of power	**Discipline (systemic)**	**Influence (episodic)**	**Force (episodic)**	**Domination (systemic)**
Breda	✓	✓		
Eindhoven			✓	✓
Emmen				✓
Groningen		✓	✓	

Source: the author

Recognising and managing the tension between exploration and exploitation is a central requirement in a theory of organisational learning. This tension is seen in the episodic power to feed-forward learning processes (exploration) and the systemic power to feedback learning processes (exploitation). In the municipalities of Breda and Eindhoven it can be observed that learning processes interact with knowledge exploitation and exploration dynamics throughout the analysis levels. It can be concluded from the empirical data that their benchlearning strategies are based on a balance between episodic and systemic modes of power. At the same time a difference can be observed: the municipality of Breda relies on forms of power affecting costs and benefits of behaviour and the municipality of Eindhoven relies on forms of power restricting available behaviours.

Emmen lacks exploration of the benchmark information since episodic power is inadequately nurtured. In the municipality of Groningen it can be observed that the episodic power is disconnected from its systemic forms of power. Systemic forms of power that might institutionalise new ideas and engender new, consistent innovations are not employed. From the empirical data it can be observed that new ideas are not integrated in group-level activities.

So, is there a relationship between the strategic stances adopted and the forms of power identified in the four cases? How do the strategic stances relate to feed forward and feedback mechanisms of benchlearning? Some patterns emerged by comparing the processes of benchlearning occurring in the four municipalities with the associated types of strategic stances.

Table 6.3 Associated forms of strategic stances within the case municipalities

Associated form of strategic stance	Defender	Reactor	Analyser	Prospector
Breda			✓	
Eindhoven				✓
Emmen		✓		
Groningen	✓			

Source: the author

As could be observed in table 6.2, in both Breda and Eindhoven learning processes of intuition-interpretation-incorporation-institutionalisation interact with knowledge exploitation and exploration dynamics throughout all analysis levels. These two case studies suggest that benchlearning is positively associated with an analyser and a prospector stance. The case studies suggest that the analyser stance (Breda) relies on forms of power affecting costs and benefits of behaviour (discipline and influence) and that the prospector stance (Eindhoven) relies on forms of power restricting available behaviours (force and domination). Both strategic stances balance exploitation (feedback) and exploration (feed forward) dynamics. The empirical data of this research study suggest that the reactor (Emmen) and defender stance (Groningen) do not balance exploitation and exploration dynamics. The reactor stance relies on a systemic mode of power which is domination. The defender stance in this research study relies on episodic forms of power: influence and force. Both Emmen and Groningen do not succeed in moving learning to the level of institutionalisation. The analysis of the four case studies suggests that high levels of benchlearning will be facilitated by a political strategy that balances exploitation and exploration dynamics. This balance appears to be more important than the forms of power chosen (forms affecting costs and benefits of behaviour or restricting available behaviours) for benchlearning.

6.6 Conclusion

In this chapter, it has been examined which forms of strategic stances and power can be seen in the four cases and how these different forms affect the learning outcomes of benchmarking. The chapter relied on two classifications: (i) the typological classification of Miles et al (1978) towards strategic stances; and (ii) the conceptualisation and typology of power as formulated by Lawrence *et al*. (2005, 2012) and Lawrence and Robinson (2007), connecting learning processes to socio-political processes.

This chapter offered insight into how strategic stances relate to feed forward and feedback mechanisms of benchlearning. The study finds empirical support for that learning outcomes from benchmarking are positively associated with an analyser and a prospector stance and negatively with a defender stance. This suggests that organisations that employ an analyser or prospector stance open the process of strategy making up to the notion of benchlearning. By contrast, municipalities that adopt a defender stance are likely to face difficulties to integrate the personally constructed cognitive maps and integrating them in a way that develops a shared understanding among the group members. Municipalities that adopt a reactor stance are facing difficulties to move the benchlearning from the group to the organisation level.

The empirical data suggest that the analyser stance relies on forms of power affecting costs and benefits of behaviour and that the prospector stance relies on forms of power restricting available behaviours. At the same time it can be concluded that a balance between exploitation (feedback) and exploration (feed forward) dynamics, as can be seen in Breda and Eindhoven, appears to be more important than the forms of power chosen to support benchlearning. The importance of a political strategy that balances exploitation (single-loop learning) and exploration (double-loop learning) dynamics means that municipalities in which this balance is not present experience limited benchlearning. Mintzberg and Waters (1985) refer to such a balance as the directing / responding dialectic (p. 272).

The central contention of this chapter is that prospectors and analysers reach a different learning outcome than defenders and reactors. For a strategic change process to take place, double-loop learning is required. To support a strategic change process resulting from double-loop learning, a balance between episodic and systemic forms of power is suggested. Whereas the first group is willing to change policies when it is perceived as necessary, the strategic stances of the second group result in narrow process improvements (single-loop learning).

7. Performance in the eye of the beholder

"Outside of any specific context, performance can be associated with a range of actions from the simple and mundane act of opening a car door, to the staging of an elaborate re-enactment of the Broadway musical 'Chicago'. In all these forms, performance stands in distinction from mere 'behaviour' in implying some degree of intent".

(Dubnick, 2005, p. 391)

7.1 Introduction

Nobody disagrees with the need for performance improvement, but how performance improvement is defined, and seen to be provided, is far from resolved. This chapter explores the way in which powerful stakeholders in the building and housing sector understand performance. It is assumed that persons at different positions in the municipalities vary in terms of the pressures they experience most by these stakeholders, which influences their definition of performance. The *content* of benchlearning in this thesis refers to the influence of different definitions of performance used by various institutional actors, allowing for a focus on the quality of achievement and/or the quality of actions. This perspective is phrased by Pettigrew as follows (1987, p. 649): "*Where we sit not only influences where we stand, but also what we see*".

To have a better insight in the way in which powerful stakeholders in the building and housing sector understand performance and how this affects the perceived dimensions of performance in the case studies, the following research questions are explored in this chapter:

- Who are the main groups of stakeholders in the building and housing sector influencing municipalities and how do they understand performance?
- What is the influence of different definitions of performance by various institutional actors on benchlearning?

The institutional context of the Building and Housing Departments is polycentric with multiple stakeholders. There are overlaps, connections and tensions between these multiple stakeholders. It is argued in this chapter that the intention to use performance information derived from municipal benchmarking for learning purposes is difficult to sustain due to the different definitions of performance as used by the various stakeholders. Initially, a content analysis of interview transcripts was undertaken by scanning for direct and indirect references to performance and related concepts, such as quality, then aggregating all verbatim statements. A second step of analysis was used to capture the process of talking about performance. Longer extracts of the interviews were selected in which interviewees were threading their way through an understanding of performance. A document analysis was done to identify the view towards performance as described in the municipal policy documents. The outcomes from these steps allowed for the identification of the definition used per stakeholder.

The chapter begins by reviewing definitions of performance in section 7.2. The powerful stakeholders who can attempt to shape or steer the performance of the Building and Housing Departments of municipalities have been mapped by use of the Performance Regime Model as developed by Talbot (2008) – see section 7.3. In section 7.4 it has been attempted to plot the understanding of different functional disciplines towards performance within each case municipality. Section 7.5 gives insight in how performance is defined in written policy documents. Some concluding observations are made in the last section of the chapter regarding the influence of different definitions of performance by various institutional actors on benchlearning (7.6).

7.2 Defining performance

Several authors (e.g., Boyne and Gould-Williams, 2003; Leeuw and Van Thiel, 2002; Moynihan and Pandey, 2005) observe that performance is a construct that has no necessary and sufficient operational definition. The conventional definition of performance uses the metaphor of the production process. Performances are the outputs and outcomes of a process. An alternative view sees performance as the realisation of public values.

Performance as a production process is currently the dominant perspective on public sector performance. In this perspective public administration is an open system which converts inputs into outputs. The model has been redefined over the years covering the whole chain from input to outcome. The outcomes of public services are either collective, or consist of externalities that are not taken into account by individual customers. Outcomes can be intermediate or final. The final outcomes in particular are influenced by the environment (Van Dooren *et al*., 2010, p. 21). A growing awareness of the inadequacies of this simple model for the public sector led several public administration scholars to redefine the model (e.g., Hatry, 1999; Poister, 2003; Pollitt and Bouckaert, 2004; Van Dooren, 2006). The metaphor of performance as a production process leaves the question unanswered what the defining characteristics of performance are.

The alternative view, *performance as the realisation of public values,* suggests that the lacuna defining the characteristics of performance may be filled by the literature on public sector values (Magd and Curry, 2003; Van Dooren *et al*., 2010). In this view, performance can be made operational as realizing public values. Several studies attempted to sort out the concept of public value (e.g., Moore, 1995; Jorgensen and Bozeman, 2002 and 2007). Public values represent a psychological and sociologic construct referring to values that persons hold, independently of the production of goods and services. So, public values exist independently of production processes for outcomes. Based on empirical evidence and theoretical input, Andersen *et al*. (2012) propose that public values can be classified in seven dimensions relating to four forms of governance (hierarchy, clan, network and market). The seven public value dimensions are called 'the public at large', rule abidance, societal interests, budget keeping, efficient supply, professionalism, and user focus.

To summarise, performance as a production process and performance as the realisation of public values ask different questions about the same issue – the public interest. A performance assessment will analyse to what extent public organisations and programs further the public interest. A value assessment will ask questions about the values that prevail, whether they are in conflict or whether they complement (Van Dooren *et al.*, 2010, p. 24).

Dubnick (2005, p. 25) adds four dimensions of performance to the debate, integrating both views towards performance. The dimensions proposed by the author cover the whole chain from input to outcome and as well contain a value judgment. He indicates that performance is about intentional behaviour, which can be individual or organisational. He refers to two aspects of performance: (a) the quality of the actions being performed, and (b) the quality of what has been achieved because of those actions. The attention given to the significance of the two perspectives can be high or low. This allows distinguishing between four dimensions of performance (see table 7.1). The view towards performance as suggested by Dubnick will be used for the sake of analysis in this chapter.

Table 7.1 Four dimensions of how performance is understood

		Focus on quality of achievement	
		Low	**High**
Focus on quality of actions	**Low**	Performance as production (P1)	Performance as good results (P3)
	High	Performance as competence / capacity (P2)	Performance as sustainable results (P4)

Source: based on Van Dooren (2010) and Dubnick (2005)

The most basic dimension of performance focuses the attention on the tasks being carried out by the performing agent (P1). Performance in this view is intentional behaviour of actors. As such, this conceptualisation is relatively neutral in nature, but also very broad. The other dimensions of the concept 'performance' contain a value judgment. Performance has a quality that can be either high or low.

When performance is about the quality of the actions, and not so much about the quality of the achievements, performance equals competence or capacity (P2). This dimension of performance is closely linked to the total quality management (TQM) approach that became so popular in the 1980s. TQM is a managerial perspective that puts forward the *capacity to perform* as a surrogate for results. Under the assumption that a highly competent performer will be more likely to generate more and better results from an activity most of the time, performance becomes associated with the competence of the performing institution (Dubnick, 2005, p. 392). High performing public sector organisations are organisations that have superior capacity. Performance in this sense is also related to constitutional competence: public administrators must be competent in their knowledge of institutional and legal responsibilities. The measurement of the competencies, skills, experience and knowledge of the performer becomes pivotal under this perspective, and performance improvement translates into strategies for enhancing the capacities, skills, etc. of the performers.

When performance is about the quality of the achievements and not so much about the quality of the actions, performance equals results (P3). The capacity of the organisation is not the focus of this conceptualisation. In this case, it is the results that count. Van Dooren (2006) argues that results may be both the outputs and the outcomes of the public sector (p. 15). This view of performance is the most prevalent type in the literature on performance associated with new public management. Dubnick (2005, p. 27) observes that results performance is typically discussed in quantitative terms.

When performance is conceptualised with the attention for both the actions and the achievements of the organisation, it may be typified as sustainable results (P4). Performance refers to the productive organisation, i.e. an organisation that has the capacity to perform and converts this capacity into results – outputs and outcome (Van Dooren *et al*., 2010, p. 3). From this perspective, performance is comprised of those actions that shape or determine the different factors in the production function. This can include decisions or acts regarding the mix of inputs, how they will be processed, what technologies will be used, where and when the production occurs, the disposition of outputs, etc. (Gold 1965, in: Dubnick, 2005, p. 27).

7.3 How is performance defined by institutional stakeholders?

The Performance Regimes Model as developed by Talbot (2008) is taken as a starting point to map the main groups of stakeholders in the building and housing sector. Thereafter, insight is given into how performance is conceptualised by these stakeholders.

The term 'performance regime' is adopted by Talbot (2008) to convey a combination of the institutional context of performance steering and the nature of actual performance interventions (or in this research study: benchmarking-related actions). Talbot (2008, p. 1594) suggests that:

> "*It is the totality of a performance regime which potentially shapes or steers performance for specific organisations rather than the narrow purchaser-provider or principal-agent assumptions often made about performance drivers. Public bodies in effect have multiple principals and this is still under-explored and under-theorised*".

Talbot (2008) uses regime in the more usual political studies sense of the constellation of institutional actors and their formal and informal cooperation and competition (p. 1583). The author suggests that alongside formal powers, there are likely to be strong informal influences and practices which shape the total context within which the performance of organisations are steered. The performance regime focuses on two main issues: which *institutions* can and do have the powers to steer or to hold to account public agencies? And what sort of *interventions* do they actually make to enact those powers.

The Performance Regimes Model of Talbot (2008) identifies institutional pressures from the following groups of stakeholders:

i. Central Ministers, Ministries and Departments. They (may) have a crucial role in determining both an overall policy approach to performance throughout the public sector as well as in (potentially) setting targets for the core executive itself.

ii. Legislature. Regulation inside government means that one public organisation shapes the activities of another (excluding judicial actors); the overseer is at arm-length from the organisation being overseen; and the overseer has some kind of mandate to scrutinise the behaviour of the 'regulatee' and seek to change it (e.g. Parliament).

iii. Audit, inspection and regulatory bodies. Audit, inspection, and regulatory bodies effectively steer towards specific forms of performance in public service organisations. These may include conformance to internal process standards and to service delivery standards and even in some cases to what types of performance measures should be used.

iv. Judicial and quasi-judicial bodies, inquiries. These have the potential to shape and steer the performance priorities of public services by setting standards of best practice or criticizing poor process or substantive performance (e.g. judicial reviews, public enquiries, etc.).

v. Professional Institutes. These have an impact on organisational practices by setting standards for professional practice.

vi. Users and user organisations. Are seen as an important factor in performance design (customer service). There are several mechanisms whereby users can have a more active role: if they have choice of provider, if their user organisations have formal representational powers (including indirect influence through other institutional actors such as individual complaints to audit bodies or complaints through an Ombudsman).

vii. Partners (through contracts). Where agencies are forced to establish contractual or quasi-contractual arrangements in purchaser-provider public sector quasi markets then such contracts will usually themselves embed performance steering criteria that are relevant to performance regimes. The contracting component is not necessarily restricted to partner organisations—there are often also "vertical" contracts—but this group is about horizontal contracts or quasi-contracts between partners that provide them with leverage over performance of agencies (sometimes called "service level agreements").

The four dimensions of how performance is understood (see table 7.1) are used to identify what emphasis these powerful stakeholders place on different dimensions of performance. The mapping is presented in table 7.2 below. Institutions in the groups of stakeholders referring to judicial and quasi-judicial bodies and partners (through contracts) could not be detected.

The classification of the performance dimension per group of stakeholders has been made based on their content role and the sort of interventions they actually make to enact those powers. The Ministry of the Interior has a crucial role in determining an overall policy approach to performance throughout the public sector by setting the framework for building and housing (whereas planning of these tasks has to be worked out at the local level). The Provincial Executive has a crucial role, through inter-administrative supervision, in the monitoring of the quality criteria for the authorisation of the all-in-one permit for physical aspects (WABO) and the monitoring and enforcement of legislation for the physical environment. The quality criteria focus on the performance dimensions competence and capacity. The Inspection Living Environment and Transport promotes compliance and implementation of laws and regulations in the sector with an emphasis on good results. Several professional institutes exists setting standards for professional practice. To ensure sustainable results in the sector, a key protocol for the collective quality standards for building permits has been formulated through the involvement of these professional institutes. This is a national method for testing building plan applications to the structural laws and regulations. In addition to this, a registration system allowing the registry of construction errors that relate to structural safety can be used by persons working in the sector. Users and user organisations are seen as an important factor in customer feedback. TNS NIPO offers a Customer Satisfaction Survey to municipalities (for which they have to pay) whereby users can have a more active role to ensure sustainable results in the building and housing sector.

Table 7.2 Groups of stakeholders influencing the Building and Housing Departments

Groups of stakeholders	Institutions	Interventions	Performance dimension
Central Ministers, Ministries and Departments	Ministry of the Interior	The Ministry has a crucial role in determining an overall policy approach to performance in the building and housing sector.	P4 Performance as sustainable results
Legislature	Provincial Executive	Quality criteria for the authorisation of the all-in-one permit for physical aspects (WABO) and the monitoring and enforcement of legislation for the physical environment. Inter-administrative supervision by the Provincial Executive and horizontal monitoring of the Board of Mayor and Aldermen and the municipal council.	P2 Performance as competence / capacity
Audit, inspection and regulatory bodies	Inspection Living Environment and Transport (previous VROM Inspection)	The Inspection promotes compliance and implementation of laws and regulations for building, housing, space and environment by companies, institutions, citizens and governments. The Inspection puts forward surveys, makes recommendations, posts suggestions and investigates incidents.	P3 Performance as good results
Professional Institutes	Association BWT NVBR (Dutch Association for Fire and Disaster) PMGG (Platform environmental enforcement large municipalities) ABC	Key protocol for the collective quality standards for building permits. This is a national method for testing building plan applications to the structural laws and regulations. The project organisation (iTP, previously known as LTP and CKB) adds the received audit results together for evaluation by municipalities themselves, but also for central government. Municipalities can thus compare their results. The central government can evaluate implementation of the law on a national level. The protocol is managed by professional institutes and supported by the Ministry of Internal Affairs and the Inspectorate (for the living environment and transport). ABC is an initiative of the Platform Structural Safety. The registration system includes the registry of construction errors that relate to structural safety. The goal of this registration is to learn from past mistakes and hence ensure safer	P4 Performance as sustainable results

		buildings. Alerts can be done by persons working in the sector. After that the alerts will be made anonymous and entered into the database. By means of analysis causes and lessons are determined.	
Users and user organisations	TNS NIPO	Customer Satisfaction Survey Building and Housing Supervision. The study focuses on customer satisfaction with the services provided by the application of a light or regular building permit.	P4 Performance as sustainable results

Sources: Dubnick, 2005; Talbot, 2008; Van Dooren, 2010. Further compilation by author based on several internet resources

Table 7.2 shows that differences exist in the content of the influences from the various stakeholders and the emphasis they place on different dimensions of performance. It can be observed that the institutional context of performance in the building and housing sector is complicated and exists in a complex web of direct and indirect influences. The groups of stakeholders with a direct influence over the Building and Housing Departments (legislature and audit, inspection and regulatory bodies) address a different dimension of performance than the institutions that have indirect influence. The groups of stakeholders with a direct influence see performance as competence/capacity or good results, whereas the groups of stakeholders with an indirect influence (professional institutes and users and user organisations) see performance as sustainable results. The Ministry of the Interior focuses on sustainable results in determining an overall policy approach to performance in the building and housing sector. It is observed that dimensions of performance vary depending upon the groups of stakeholders. Thus the 'success rates' are measured against different performance dimensions: some stakeholders may focus on quality of achievement, others on quality of actions or on both.

It is expected that the perception of performance by the interviewees is shaped and steered by multiple internal and external stakeholders (the performance regime of the Building and Housing Departments). In the subsequent section insight is given into how performance is conceptualised by the internal institutional actors of the Building and Housing Departments along the lines of the dimensions of performance as referred to in table 7.1.

7.4 How is performance defined by the interviewees?

In this section it is attempted to plot the understanding of performance from different categories of functions (i.e., functional perspectives) towards performance within each case municipality. Individual actors may have changing preoccupations and priorities which may result in changing emphases in what they see as good performance (Talbot, 2008, p. 1585). Jansen (2008) concludes in his empirical study about the use of performance information in three Dutch municipalities that politicians (i.e. strategic level) have a different perspective on performance, as compared to managers (i.e. managerial level). His results show that managers have an internal perspective, whereas politicians seem to have a citizen perspective and a financial perspective on performance.

The categories of functions used for the analysis in this section is based on the three major types of managerial decision making and control activities as suggested in section 3.7, namely, strategic planning, management control and operational control. In the municipalities studied, the persons involved in strategic planning include top management and staff specialists (Alderman, Division Head). Management control is the responsibility of line managers (Section Head, Team Leader and Policy Officer). Operational control is done by first-line supervisors (Coordinator, Officer).

In table 7.3 it is attempted to plot the understanding from different functional perspectives towards performance within each case municipality (R = respondent).

Table 7.3 Functional perspectives towards performance within the case studies

		P1 (production)	**P2 (competence/ capacity)**	**P3 (good results)**	**P4 (sustainable results)**
Breda	Strategic				R1, R17 and R18
	Managerial		R11		R15 and R16
	Operational		R12	R14	R13
Eindhoven	Strategic				
	Managerial			R21, R23, R24	
	Operational				R22
Emmen	Strategic				R7
	Managerial			R2, R3 and R6	R4
	Operational			R5	
Groningen	Strategic				R10
	Managerial			R9 and R19	R8
	Operational			R20	

Source: the author

It appears from the long list that the interviewees produce a cacophony of definitions about performance. The key terms selected from the verbatim statements offer as many different definitions as interviewees. Together, the interviewees experience a concept of performance which encompasses multiple meanings. Several observations can be made from table 7.3:

- Employees within the same municipality can perceive performance improvement, the intentionally changing of an existing state of affairs into a superior state of affairs, differently.
- All references of performance found in the research contain a value judgement. According to all interviewees, performance has a quality that can be either high or low (P2, P3 and P4).
- All interviewees involved in strategic planning conceptualise performance with attention for both the actions and the achievements of the organisation (P4).
- Eight out of the twelve interviewees involved in management control view performance as being about the quality of the achievements with a low focus on the quality of the action. For them it is the results that count (P3).
- In Eindhoven, Emmen and Groningen the main focus is on the quality of the achievements (P3). In Breda, the key focus is on both the actions and the achievements of the organisation (P4).
- Only interviewees in Breda make reference to performance as capacity and competence (P2). In the other three municipalities, reference is made to either performance as good results (P3) or sustainable results (P4).
- First-line supervisors in Breda show a wide variance of perceptions of performance, although all agree that performance contains a value judgment (a quality that can be either high or low).

From the empirical data it can be observed that there are perceptual differences between employees working for the same municipality. It appears that a difference in perspective between top management (i.e. strategic level) and line managers (i.e. managerial level) is an important reason for differing perceptions. Whereas top management conceptualises performance with attention for both the actions and the achievements of the organisation (P4), the majority of the line managers view performance as being about the quality of the achievements and not so much about the quality of the action. For the majority of the line managers, it is the results that count (P3). Results performance is typically discussed in quantitative terms.

It could be observed from table 7.2 that the groups of stakeholders with a direct influence over the Building and Housing Departments (legislature and audit, inspection and regulatory bodies) address a different dimension of performance than the institutions that have indirect influence. The groups of stakeholders with a direct influence see performance as competence/capacity (P2) or good results (P3), whereas the groups of stakeholders with an indirect influence (professional institutes and users and user organisations) see performance as sustainable results (P4). Line managers interact directly with audit companies, inspection and regulatory bodies in the building and housing sector. It can be observed that the functional group perspective of performance among line managers as well sees performance as good results (P3). In the same line of reasoning it can be observed that groups of stakeholders with an indirect influence see performance as sustainable results (P4), which is in agreement with the conceptualisation of top managers. It is suggested that the interaction with external stakeholders affects the functional group perspective of performance. Another possible explanation for the difference in focus between the strategic and the managerial level can be found in the purpose to which performance might contribute. Performance information use at the managerial level seems to be related to internal management, whereas the strategic level is focusing on how to improve policy or management (see table 1.2).

The influence of different definitions of performance by various institutional actors can be further explored by looking into the policy documents of the municipalities. Policy documents tend to be compiled by (or at least are approved by) top managers (i.e.

strategic level). In section 7.5 it is explored if the functional perspectives towards performance can be seen as an operationalisation of the goals described in the policy documents, or if they reflect a certain degree of decoupling.

7.5 How is performance defined in policy documents?

In this section a brief reference is made to descriptions of performance in written policy documents of a particular municipality.

In the Construction Policy (2008-2011) of the city of Breda, the ambition is formulated to obtain transparency in the process and to enlarge the quality of the work by giving a clear insight into the priorities and capacity allocation within the sector. In the document, the key dimensions of performance (related to the building and housing sector) that are strived for are reflected:

> "*We give our customers the best that we can offer - for themselves - and for the people in their environment. We do this through our products and by taking customer-centred decisions. We strive to distinguish ourselves by offering the best products to our customers and to offer them the best quality at an acceptable price*" (p.14).

The Housing Policy 2007-2010 of the municipality of Breda, raises the question of how the quality of new and existing housing can be brought to the expected population growth in 2020. The perspective is that the demands of citizens are changing. It is a demand for more quality and comfort, but also for identity and small-scale development. The Residential Vision 2007-2010 of the municipality includes four central themes: vitality, quality, choice and 'doing together'. The quality of housing is dependent on various factors. In the field of housing, the engineering, technical and environmental quality is said to be important (p.33). As guiding quality documents the department is using the "Guide on Demolition" and the action plan and policy guidelines on structural safety. The key protocol for the collective quality standards for building permits was introduced by the department in 2008. In the written documents, a high focus can be observed on both the quality of achievement and the quality of actions (P4). In addition to this, attention is paid about the values that prevail, such as user focus and societal interests.

In the strategic plan of Eindhoven for the years 2006-2009 the mission of the Building and Housing Department is described: "*To promote the quality of the current and future built environment in Eindhoven through education, assessment, supervision and enforcement. By quality we mean: health, safety, durability, aesthetics, cultural history, organisation, and use*". One of the multi-annual strategic goals is to rank among the top five of the Netherlands, related to the benchmark BWT.

The service delivery of Eindhoven is standardised in a Quality Charter that is revised and supplemented on an annual basis together with the citizens. In the clients monitor for construction and other permits, the results of the investigation into the quality of the services are reflected. The monitor makes it possible to identify points for improvement. The total score is dependent on four indicators: quality of contact and information provision; quality of the testing phase; quality of the monitoring phase; and quality of obtaining the license. In the written documents, a high focus can be observed on both the quality of achievement and the quality of actions (P4). In addition to this, attention is paid about the values that prevail, such as professionalism and societal interests.

The Living Quality Charter of Emmen aims to ensure the quality of living and where necessary to further strengthen this quality. The focus on housing quality occurs through both concrete quality themes and a focus on process issues, wherein the relationship between government, housing consumers and providers are central. These relationships should be designed to promote the quality of living. Attention should be given to consultative structures, information and demand-oriented services, formal agreements about living and quality of life and realisation. The quality themes are further detailed in various policy programs and notes. The Charter is the local implementation of the Agenda for a higher quality of living in the north of the Netherlands. The three Northern provinces (Drenthe, Friesland and Groningen) work together in a partnership (SNN). The objective of this partnership is to reinforce the economic position of the Northern Netherlands. In the Charter a number of quality themes have been appointed which are considered important for the further elaboration of the quality of life. These are among others: equilibrium on the housing market; differentiation of the living environment; accessibility; sustainable building; and environmental quality. In the written documents, a high focus can be observed on both the quality of achievement and the quality of actions (P4). In addition to this, attention is paid about the values that prevail, such as societal interests.

In the annual plan (2009) of the municipality of Groningen three out of five targets are related to a quality dimension. One target is related to the quality enhancement of the building code review and the other two targets aim the quality enhancement of the inspection of building permits. The enforcement policy of the municipality should be seen as a policy framework within which a number of assumptions, constraints and priorities have been formulated. Success is achieved when the Building and Housing Department carries out a recognisable and transparent enforcement policy for citizens, businesses and institutions. The key dimensions of performance that are strived for include a qualitative approach towards licensing, investigation and enforcement. In the written documents, a high focus can be observed on both the quality of achievement and the quality of actions (P4). In addition to this, attention is paid about the values that prevail, such as 'the public at large' and professionalism.

It can be observed that the policy documents reviewed mainly refer to performance as sustainable result. In addition to this, it can be observed that the policy documents focus on two questions: (i) to what extent the municipalities and their programs further the general interest and (ii) whether they are in conflict or whether they complement each other. As such, performance as defined in these policy documents is construed around both actions and achievements.

7.6 Conclusion

The field evidence establishes that the intention to use performance information derived from benchmarking for learning related purposes is difficult to sustain due to the insufficient clarity regarding the influence of different political strategies, and perspectives on performance as given by its context. The analysis in this chapter suggests that there are perceptual differences among the various stakeholders in the building and housing sector about the dimension of performance to strive for.

There is a roughly even split between (external) stakeholders emphasising competence/capacity or good results and stakeholders emphasising sustainable results, with no stakeholder emphasising performance as production. This finding is reinforced by table 7.3 where two respondents are classed as P2 (competence/capacity), 11 respondents as P3 (good results) and 11 respondents as P4 (sustainable results). Roughly speaking, respondents at the strategic level are concerned with a high quality of both the actions and the achievement (sustainable results), whereas those at managerial level are mainly concerned with a high quality of achievement (good results).

At the level of interviewees differences could be detected among the cases in how dimensions of performance are referred to. In Breda, the key focus of top management and line managers is on both the actions and the achievements of the organisation (P4, sustainable results). Within the other three municipalities, the main focus is on the quality of achievements (P3).

In table 7.2 the empirical evidence suggests that the institutions with a direct influence over the Building and Housing Department address a different dimension of performance than the institutions that have indirect influence. The groups of stakeholders with direct influence see performance as competence/capacity and good results whereas the groups of stakeholders with indirect influence see performance as sustainable results. This research study suggests that the differences in the perspective on performance between the stakeholders are the consequence of different levels of influence (and hence possibilities of intervention) on the side of the stakeholder.

In general, it can be observed that the policy documents both refer to performance in terms of sustainable results. The focus on sustainable results is unsurprising as such documents will be compiled by those at the senior strategic level of the case organisations. In practice, performance in the municipalities is construed more broadly within policy documents than in the statements made by the stakeholders and the interviewees.

It can be concluded that the challenge to define performance is not to formulate one definition or model that attempts to account for all possible variables. According to this research, performance is subjectively constructed and differs per person: performance is in the eye of the beholder. Being aware of the different dimensions strived for by internal and external stakeholders can support that the 'success rates' are measured against relevant performance dimensions.

So, in what way is the use of performance information derived from benchmarking influenced by different definitions of performance? In chapter 5 it was concluded that a balance between exploitation (feedback) and exploration (feed forward) dynamics exists in both Breda and Eindhoven. A hybrid approach could be observed in these municipal organisations combining interpretative approaches at group level with a mixture of hierarchical and network strategies to move learning from one analysis level to the other. The difference in the main focus towards performance as could be seen in Breda and Eindhoven, respectively with a main focus on sustainable results and good result, do not seem to influence the level of benchlearning. A possible explanation for the success of these municipalities can be found in the combination of interpretative approaches at group level with a mixture of hierarchical and network strategies. The chosen strategy brings together different insights (at group level and within the network), allowing for an exchange on perceptual differences, facilitating the formulation of a preferred future scenario for performance improvement.

8. Conclusions

8.1 Introduction: thesis research objective and research questions

The research reported in this thesis focuses on benchlearning within municipalities and between groups of municipalities with a specific emphasis on the Dutch building and housing sector. Contributions from a number of different fields were reviewed in order to ground the research in the literature. Two theories – organisational learning and organisational change – were utilised to frame the theoretical research questions and interpret the empirical findings. The literature as presented in the literature review depicts that existing studies on benchmarking and performance management do not search for underlying classifications and mechanisms related to the different uses of performance information. It is suggested in the literature that a performance measurement and management framework should include an examination of the underlying processes and the inner context of the organisation to support its quality. Therefore, this thesis explicitly addressed the role of the context in which process and content are embedded as suggested by Pettigrew (1987), with an emphasis on how organisational learning informs and extends the understanding of benchlearning. This led to a focus on the following three research themes:

1. Process: the interconnections between organisational learning and the incorporation and use of benchmark information;
2. Context: the influence of different forms of strategic stances and power on benchlearning and improved performance;
3. Content: the influence of different definitions of performance by various institutional actors on the use of performance information from benchmarking.

The first research theme has been supported by using the "4i framework of organisational learning" as presented by Crossan *et al*. (1999). The second research theme relied on two classifications: (i) the typological classification of Miles *et al*. (1978) towards strategic stances; and (ii) the conceptualisation and typology of power as formulated by Lawrence *et al*. (2005, 2012) and Lawrence and Robinson (2007), connecting learning processes to socio-political processes. The third research theme has been supported by using the Performance Regime Model as developed by Talbot (2008).

It is suggested in this research study that the relative success of different approaches to dimensions of change and organisational learning may be a key explanatory factor for differences in the success of benchlearning.

In order to address the research themes, four case studies at municipal level (Breda, Eindhoven, Emmen and Groningen) were carried out. Unstructured interviews (pilot interviews), document analysis and semi-structured interviews were used to collect the data.

The findings of this research study are summarised in section 8.2. Next, section 8.3 discusses the implications of this thesis's findings for theory, methodology and method, and policy and practice. Section 8.4 describes the limitations of this research study and section 8.5 identifies viable directions in which future research might go.

8.2 Summary of findings

This research study has resulted in several conclusions and suggestions that can lead to an increased understanding of municipal benchmarking in relation to its process, context and content.

Municipal benchmarking and organisational learning (process)

The cases provided some insights about the interconnections between organisational learning and the incorporation and use of benchmarking information. The first research question aims to explain which levels of analysis, learning processes, learning barriers and supporting factors can be observed in municipal benchmarking. In this research study, it is observed that benchlearning takes place in two parallel communities: within a municipality and between groups of municipalities. The 4i framework of organisational learning required the inclusion of the inter-organisational level, to provide a helpful starting point for studying municipal benchmarking.

So what empirical observations can be made when applying the adjusted learning framework to the four case studies? The research study found that the process of benchlearning involves a number of underlying classifications in which learning is carried within a municipality and between groups of municipalities. The intended uses of benchmark information set the scene for the carriage of benchlearning through the different levels. In the municipality of Groningen, for example, where benchmark information is used for the sake of comparison and defensive use, it can be observed that the information is not carried from the individual to the group level. The benchmarking report is not shared within the department and not discussed within the management team. Structural-organisational learning barriers such as a lack of staff resources and a high workload of the management team and/or staff members are provided as explanations. March (1991) observes in this respect that both exploration and exploitation are essential for organisations, but that they compete for scarce resources.

Case study findings illustrated that shifting from individual learning to learning among individuals or groups appeared difficult in the municipalities where interpretative processes at the group level within the municipal organisation were absent. In the more successful case studies (Breda and Eindhoven), the emphasis appeared to be more on a combination of interpretative approaches at the group level with a mixture of hierarchical and network strategies. Taking the case findings of the municipality of Breda as an example, it can be observed that the head of the department takes a leading role in the feed forward of the benchmarking information and is actively involved in external networks such as expert groups and the Association of BWT. This observation suggests that actions taken within the municipalities are not as clearly separated from actions taken within the network. In addition to this it can be observed that the situation in practice is more complicated than initially suggested. Interaction between municipalities at the inter-organisational level is positively influencing benchlearning, which is not restricted to the benchmark network itself. So, the case material suggests that interaction among professionals at the inter-organisational level is a factor that positively influences benchlearning within the municipality.

It can be observed from the empirical data presented that the barriers to benchlearning are not necessarily a decisive reason for limited learning across all analysis levels. Stimulating supporting factors can be seen as a more important strategy to stimulate benchlearning than pulling down barriers.

The research study suggests that supporting factors for benchlearning should be predominantly looked for at the individual, group and organisational level instead of the inter-organisational level. From the analysis it can be observed, that the awareness of which factors support benchlearning (and hence the possibility to steer on these), has a strong impact on moving the learning from one analysis level to the other within the municipality. The interviewees of both Emmen and Groningen showed very limited reference to supporting factors. The empirical data from Breda and Eindhoven suggest that in order to facilitate benchlearning intrinsic motivation and communication skills matter, supported by a high level of cooperation (team work), a flat organisational structure and interactions between individuals. The participation in external networks and a continuous search for dialogue appear to have a positive influence on benchlearning as well. This suggests that in order to facilitate benchlearning more attention should be paid to the quality of relationships, conversations and interactions taking place between groups of municipalities and within municipalities.

Municipal benchmarking and power and politics (context)

The cases provided some insights about the interconnections between municipal benchmarking and power and politics. The research questions aimed to explain the influence of different forms of strategic stances and power on benchlearning and improved performance. Research findings suggest that an unbalanced use of both episodic (emergent) and systemic (deliberate) forms of power can be signalled as a determined cause of the interrupted knowledge exploitation and exploration dynamics within a municipality. The importance of a strategy that balances exploitation and exploration dynamics means that municipalities in which this balance is not present experience limited benchlearning. To take the municipality of Groningen as an example, it can be observed that the episodic politics are disconnected from its systemic forms of power. Systemic forms of power that might institutionalise new ideas and engender new, consistent innovations are not employed. From the empirical data it can be observed that conclusions as presented in the benchmark report are not integrated in group-level activities. The benchmarking report is not being used as a steering instrument or as an input to annual plans or other departmental items.

The two municipalities being more successful in carrying benchlearning through the organisation rely on different strategies. The municipality of Breda uses a strategy that is based on affecting costs and benefits of behaviour (discipline and influence) and the municipality of Eindhoven uses a strategy that is based on restricting available behaviours (force and domination). The affection of costs and benefits can be observed in terms of socialisation, training opportunities and team-based work. The restriction of costs and benefits can be observed for example in terms of the mal acceptance of the benchmark results by the alderman.

It is suggested in this research study that high levels of benchlearning will be facilitated by a strategy that is based on a balanced use of both episodic (emergent) and systemic (deliberate) forms of power and that the approach taken to this (either affecting or restricting behaviours) is not necessarily influencing this. The importance of using forms of power that balance exploitation and exploration dynamics means that municipalities in which this balance is not present experience limited benchlearning.

The empirical data suggest that benchlearning is positively associated with an analyser and a prospector stance. In addition to this, data suggest that the analyser stance (Breda) relies on forms of power affecting costs and benefits of behaviour (discipline and influence) and that the prospector stance (Eindhoven) relies on forms of power restricting available behaviours (force and domination). The empirical data suggest that the reactor (Emmen) and defender stance (Groningen) do not balance exploitation and exploration dynamics. The reactor stance relies on a systemic mode of power which is domination.

The defender stance in this research study relies on episodic forms of power: influence and force. Both Emmen and Groningen do not succeed in moving learning from one analysis level to another. Whereas analysers and prospectors are willing to change policies when it is perceived as necessary, the strategic stances of defenders and reactors result in narrow process improvements (single-loop learning). This suggests that high levels of benchlearning will be facilitated by an analyser or prospector stance.

Municipal benchmarking and dimensions of performance (content)

The cases provided empirical insights about municipal benchmarking and dimensions of performance. The last two research questions of this thesis intended to assess how the main groups of stakeholders understand performance, and how different perceptions towards performance by various institutional actors influence the use of performance' information.

Research findings suggest that to increase the quality of achievements and actions in the building and housing sector, it is important to understand how performance is constructed by, and extracted from, those who intend to improve the quality of service delivery within the municipalities. Various dimensions of performance are adopted by the powerful stakeholders in the building and housing sector, determined by the different levels of influence of the stakeholder. The field evidence establishes that the intention to use performance information derived from benchmarking for learning related purposes is difficult to sustain due to the insufficient clarity regarding the influence of different political strategies, and perspectives on performance as given by its context.

Considering the various dimensions of performance that can be adopted, the empirical data show that the focus towards performance is influenced by function. Respondents at the strategic level are concerned with sustainable results whereas those at managerial level are mainly concerned with good results. A possible explanation for the difference in focus between the strategic and the managerial level can be found in the purpose to which performance might contribute. Performance information use at managerial level is related to internal management and accountability purposes, whereas the strategic level is focusing on how to improve policy or management. It can be observed that the institutional context of performance in the building and housing sector is complicated and exists in a complex web of direct and indirect influences. In addition, it can be observed that the groups of stakeholders with a direct influence over the Building and Housing Department (legislature and audit, inspection and regulatory bodies) address a different dimension of performance than the institutions that have an indirect influence. The groups of stakeholders with a direct influence see performance as competence/capacity and good results whereas the groups of stakeholders with indirect influence (professional institutes and users and user organisations) see performance as sustainable results. The Ministry of the Interior focuses on sustainable results in determining an overall policy approach to performance in the building and housing sector.

Moreover, the empirical data suggest that, in practice, performance in the municipalities is construed more broadly within policy documents than in the statements made by the interviewees. Whereas the individuals in the organisation see performance as a production process and consequently relate performance to the quality of actions and/or the quality of achievements, the organisational perception of performance (as written in the policy documents) is also based on public value dimensions. In other words, even if the head of the Building and Housing Department has a clear understanding of the ambitions, goals and functions of participating in the benchmark, this does not necessarily mean that perspectives on performance as given by the internal and external context of the municipality are similar. Being aware of the different dimensions strived for by internal and external stakeholders can support that the 'success rates' are measured against relevant performance dimensions. The difference in the main focus towards performance as could be seen in Breda and Eindhoven, respectively with a main focus on sustainable results and good result, do not seem to influence the level of benchlearning.

The case studies showed that inner and outer contextual factors play an important part in the development of measurement systems, and in establishing a need to use approaches such as benchmarking and in encouraging commitment to their use. At the same time, the studies as reviewed in chapter two indicated that authors are divided as to which element(s) will improve performance: competition (compulsory), cooperation (voluntary) or a combination of both. This research suggests that performance improvement is influenced by functional perceptions towards performance, and that these perceptions ultimately influence the elements adopted. This is a more complex view towards defining performance, suggesting a context-dependent definition of performance.

A particular lack of clarity was found in relation to the various analysis levels and dimensions of learning that were adopted in the case studies. To provide a helpful starting point, the 4i framework of organisational learning required adjustment for studying municipal benchmarking, adding the inter-organisational analysis level. This research shows that efforts aimed at benchlearning and ultimately improved service delivery, should be directed to a multi-level and multi-dimensional approach addressing the context, content and process of dimensions of change and organisational learning.

8.3 Implications

The researcher believes that, by creating knowledge that is grounded in theory and relevant to practice, it is possible to improve the understanding of benchlearning in the public sector. The objective of this study is to enhance the theory and practice of municipal benchmarking, to impact on policy-making and to improve the ways in which approaches to dimensions of change and organisational learning are formulated and used, hence having a positive effect on the services delivered.

From a personal perspective, I see several opportunities to take the findings further and enhance the theory and practice of municipal benchmarking in my own consulting and academic research activities. My intention is to take this work further by actively sharing the findings at municipal level (national as well international) and adopting the directions derived from this research study in consulting assignments that relate to this topic. The active remaining contacts with all organisations mentioned in this study allow for ample follow up in the form of presentations, discussions and more practice oriented publications. The comprehensive literature review suggests that work on updating the literature review might lead to publication of a literature survey in an appropriate journal.

The empirical chapters as well give ample scope for publication. Among others a publication is in preparation as a joint initiative with my supervisors addressing benchmarking as a network-based form of governance in the public sector.

The following sections describe the main implications this research has on a number of levels. From a theoretical viewpoint, contributions are made to the field of PMM in general and municipal benchmarking in particular and to the empirical use of organisational change and learning theory. In terms of methodology and method, the use of case studies (including the data triangulation of interviews and documents) led to the identification of several context-specific tendencies. Finally, this research also has important implications for both policy makers and practitioners and for the design and use of municipal benchmarking in the public sector.

8.3.1 Implications for theory

This research considers several issues (such as, the interconnections between organisational learning and the incorporation and use of performance information, and the influence of different forms of power and definitions of performance) on benchlearning. From a theoretical viewpoint, concepts from the organisational change and organisational learning theory are related to the literature on benchmarking in the public sector. In particular, the interconnections between learning and change are addressed both theoretically and empirically, with the expectation for significant cross-fertilisation of ideas from these schools.

The "context, content, process model for successful organisational change" presented by Pettigrew (1987) allowed for a relevant bridge between organisational change and organisational learning and assisted in the search for underlying classifications and mechanisms related to the different uses of performance information. Analysis of the inner context of benchlearning (referring to forms of strategic stances and power used) helps to explain why some municipalities are more successful at learning than others. The content of benchlearning refers to the influence of different definitions of performance used by various institutional actors, allowing for a focus on the quality of achievement and/or the quality of actions. The process of benchlearning refers to the interaction of knowledge exploitation and exploration dynamics throughout four analysis levels (i.e. inter-organisational, individual, group and organisational levels). The inter-organisational level refers not only to the benchmark network, but to a broad arena of inter-organisational groups (groups of municipalities, professional groups etc.).

The adoption of the 'theoretical lenses' enables the investigation of phenomena taking place both within a municipality (intra) and between groups of municipalities (inter). In particular, it was possible to consider aspects related to learning processes, forms of power, strategic stances and conceptual dimensions of performance. Therefore it is suggested to draw from both organisational change and organisational learning theory to gain a deeper understanding of benchlearning in the public sector, taking the role of context, content and process as a point of departure.

8.3.2 Implications for methodology and method

The use of case studies enables the identification and examination of issues that are relevant to both theory and practice. Conducting case-based research allowed for the evolution of the research questions over time. The semi-structured interviews triangulated with a wide range of documents enabled the study of a number of issues (such as the performance dimension used) that would not have emerged if the research had relied solely on one method. Moreover, the selection of interviewees based on three major types of managerial decision making and control activities led to the identification of quite different perceptions towards performance dimensions. The literature review was useful to explore the concepts and the identification of key themes. By categorizing the research questions of earlier studies, it was possible to position the research vis-à-vis previous research efforts. The way forward suggested by the researcher to further explore benchlearning is to select a research strategy that provides for the identification of the role of context, content and process in benchlearning.

8.3.3 Implications for policy and practice

From a personal perspective, the researcher wants her work to contribute to the practice of benchlearning in the public sector. Therefore, while carrying out the fieldwork, the researcher adopted a critical perspective on the local government's actual PMM strategy. It is believed that the reflections presented in this dissertation could foster a fruitful debate on the ways in which municipal benchmarking is currently used in the public sector.

One of the main implications for policy makers stems from the complexity of the links existing between learning from performance information at the inter-organisational level and within the organisation. In particular, central government bodies, professional organisations and any other organisations that influence the learning derived from municipal benchmarking should create space to allow for interaction and exchange among policy makers and professionals. This joint up approach to municipal benchmarking includes horizontal and vertical learning. A professional body (for example the Association of BWT) appears to have a positive influence on the integration and institutionalisation of benchmark information.

One of the main implications for practitioners stems from the analysis of the success of different approaches to dimensions of change and organisational learning. First of all, the stimulation and creation of opportunities for interaction at the inter-organisational level among professionals seem to foster benchlearning. This for example implies the creation of sub-thematic learning groups at the inter-organisational level addressing different aspects of the building and housing sector. Second, the empirical data suggest that a focus on structural-relational supporting factors such as the stimulation of departmental communication and cooperation, team work and shorter hierarchical lines could lead to improved benchlearning. It is suggested that managers give priority to the strengthening of supporting factors instead of focusing on pulling down barriers. Thirdly, the empirical data suggest that in order to facilitate benchlearning more attention should be paid to the quality of relationships, conversations and interactions taking place between groups of municipalities and within municipalities. Attention for communication skills at the individual level, supported by a high level of cooperation (team work) and a flat organisational structure will positively influence on benchlearning. Fourth, the empirical data suggest that managers should adopt a strategy that is based on a balanced use of episodic (emergent) and systemic (deliberate) forms of power. This would suggest including flexibility as well as stability, adopting structures that can accommodate both stable and changing domains. This could be achieved by adopting episodic forms of power such as moral suasion, negotiation, agenda setting and limiting decision alternatives. Examples of systemic forms of power a manager could adopt are: socialisation, training, team-based work and the changing of material technologies, information systems and the physical layout. A last implication for practice is the awareness of the different dimensions strived for by internal and external stakeholders. This awareness can support that the 'success rates' are measured against relevant performance dimensions. Moreover, from the point of view of the manager, it is important to recognise that differences in perceptions of performance exist. Many times these perceptions of performance are at variance with the perceptions of the majority of the performance regime. Public managers can use the performance information derived from benchmarking for improving different forms of performance. However, no single measure is appropriate for all purposes. It is suggested to clearly formulate organisational ambitions, goals and functions to enhance performance improvement from benchmarking. In this context, cross-functional teams can take advantage of these multiple perspectives to improve the incorporation and use of performance information derived from benchmarking.

8.4 Limitations

The limitations of this research are determined by the choices made on theoretical and methodological levels and by the personal characteristics of the author. From a theoretical point of view, the conjoint use of organisational learning and organisational change theory influenced the conclusions that could be drawn from this study. However, it could be argued that the adoption of any theoretical lens would determine which aspects to consider. Therefore, it would be interesting to investigate similar topics utilizing other theories and perspectives.

From a methodological point of view, given the lack of empirical studies that look at benchlearning, this research is of a predominantly exploratory and descriptive nature. The cases considered provided the researcher with sufficient data to expand the existing theories, but other cases carried out in different settings would contribute to the gaining of a deeper comprehension of the issues examined.

In this research study, only inner contextual factors have been assessed. The influence of different forms of strategic stances and power on benchlearning has been taken as an entrance point (chapter 6). The outer context (the social, economic, political and competitive environment in which the municipality operates) has only been taken into consideration in relation to the definition of performance (chapter 7). The analysis of the inner context only partially discussed the structure and culture within the organisation through which learning have to proceed. The main emphasis has been on the (inner) political context. Nevertheless a relevant observation could be made in how inner and outer contextual learning combine to create performance. The case material suggests that interaction among professionals at the inter-organisational level is a factor that positively influences benchlearning within the municipality. Cases including the outer context and inner contextual factors like structure and culture would contribute to the gaining of a deeper comprehension of the issues examined.

The review of the literature and the empirical analysis show that there is substantial commonality of issues across the (local) public sector in the Netherlands towards improvement of benchlearning. The conclusions of this research might be of interest to public sector organisations operating in other countries in which similar issues are present, e.g. unbalanced use of episodic and systemic forms of power and the adoption of various dimensions of performance.

In terms of personal characteristics, the review of the literature, although systematic and extensive, was influenced by the researcher's subjective structuring of the process and identification of the main themes. Similar reflections could be made regarding the analysis of the material gathered in the empirical phases. In this sense, it could be argued that the personal characteristics of the researcher played an important role in this study. Therefore, it would be interesting to have other research projects carried out on similar topics by academics with different backgrounds and adopting divergent perspectives.

8.5 Further research

Review of the literature and an analysis of the data have led to the identification of a number of areas that require further research. This section presents a number of avenues for further research. Seven promising lines of research appear from this research study.

The *first* promising line of research is investigation of the relation between municipal strategic stances and participation in external networks. Is the route to high levels of benchlearning indeed based on the organisational ability to mobilise knowledge from external networks? The *second* promising line of research is to explore which routines within a municipality support the institutionalisation of the outcomes of municipal benchmarking. Do the routines developed within the benchmark network (clearly formulated goals, aligned interests of stakeholders, clear responsibilities, structured benchmark cycle and regular meetings) as well apply to the municipal level? The *third* line of research suggested is investigation of the effects of a hybrid approach (combining interpretative approaches at the group level with a mixture of hierarchical and network strategies) to benchlearning. Is this proposition, based on a limited set of four case studies, holding up in broader samples? The *fourth* line of research suggested is investigation of the relation between the stimulation of supporting factors and benchlearning. Does the observation that stimulating supporting factors can be seen as a more important strategy to stimulate benchlearning than pulling down barriers hold up in broader samples? And does this observation apply to organisational learning in general? The *fifth* promising line of research is investigation of the relation between power and benchlearning. Do the various propositions as suggested in chapter 6, based on a limited set of four case studies, hold up in broader samples, for example, is the route to high levels of benchlearning indeed based on a balance between systemic and episodic forms of power? And in what way is the choice between systemic and episodic forms of power influencing organisational behaviour and learning? The *sixth* promising line of research is investigation of the relation between strategic stances and forms of power. Does the observation that the analyser stance relies on forms of power affecting costs and benefits of behaviour and the prospector stance on forms of power restricting available behaviours hold up in broader samples? The *seventh* line of research suggested is to explore how the different definitions of performance co-exist at the inter-organisational level where institutional actors with a direct and indirect influence in the building and housing sector interact.

It is believed by the researcher that a learning mind-set towards municipal benchmarking will lead to greater depth and value of municipal benchmarking in the public sector. Municipal benchmarking is an underutilised strategic asset for enhanced organisational learning and network performance in the public sector.

List of References

Adcroft, A. and R. Willis (2005), The (un)intended outcome of public sector performance measurement, *International Journal of Public Sector Management*, Vol. 18, No. 5, pp. 386-400.

Ahrens, T. (2008), A comment on Marja-Liisa Kakkuri-Knuuttila, Kari Lukka and Jaakko Kuorikoski, *Accounting, Organizations and Society*, No. 33, pp.292-297.

Ahrens, T. and C.S. Chapman (2006), Doing qualitative field research in management accounting: positioning data to contribute to theory, *Accounting, Organizations and Society*, No. 31, pp. 819-841.

Alstete, J.W. (2008), Measurement benchmarks or "real" benchmarking? An examination of current perspectives, *Benchmarking: An International Journal*, Vol. 15, No. 2, pp. 178-186.

Amaral, P. and R. Sousa (2009), Barriers to internal benchmarking initiatives: an empirical investigation, *Benchmarking: An International Journal*, Vol. 16, No. 4, pp. 523-542.

Ammons, D.N., Coe. C. and M. Lombardo (2001), Performance comparison projects in local government: participant's perspectives, *Public Administration Review*, January/February 2001, Vol. 61, No. 1, pp. 100-110.

Ammons, D.N. and W.C. Rivenbark (2008), Factors influencing the use of performance data to improve municipal services: evidence from the North Carolina benchmarking project, *Public Administration Review*, March/April 2008, pp. 304-318.

Anand, G. and R. Kodali (2008), Benchmarking the benchmarking models, *Benchmarking: An International Journal*, Vol. 15, No. 3, pp. 257-291.

Andersen, L.B., Henriksen, B. and I. Spjelkavik (2008), Benchmarking applications in public sector principal-agent relationships, *Benchmarking: An International Journal*, Vol. 15, No. 6, pp. 723-741.

Andersen, L.B., Jorgensen, T.B., Kjeldsen, A.M., Pedersen, L.M. and K. Vrangbaek (2012), Public Value Dimensions. Developing and testing a multidimensional classification, *International Journal of Public Administration*, vol. 35, no. 11, pp. 715-728.

Andrews, R., Boyne, G.A. and R.M. Walker (2006), Strategy content and organisational performance: an empirical analysis, *Public Administration Review*, January/February 2006, pp. 52-63.

Anthony, R. N. and V. Govindarajan (2006), *Management Control Systems* (12th edition), McGraw-Hill/Irwin, USA.

Argyris, C. And D.A. Schön (1978), *Organisational learning: a theory of action perspective*, Reading, MA: Addison-Wesley.

Askim, J., Johnsen, A. and Cristopherson K-A (2007), Factors behind organisational learning from benchmarking: experiences from Norwegian municipal benchmarking networks, *Journal of Public Administration Research and Theory*, Vol. 18, pp. 297-320.

Babbie, E.R. (2010), *The practice of social research*, Wadsworth Publishing, USA.

Ball, A., Bowerman, M. and S. Hawksworth (2000), Benchmarking in local government under a central government agenda, *Benchmarking: an International Journal*, Vol. 7, No. 20.

Bate, P. And Robert, G. (2002), Knowledge management and communities of practice in the private sector: lessons for modernizing the National Health Service in England and Wales, *Public Administration*, Vol. 80, pp. 643-663.

Behn, R.D. (2003), Why measure performance? Different purposes require different measures, *Public Administration Review*, September/ October 2003, Vol. 63, No. 5, pp. 586-606.

Berends, H. and I. Lammers (2010), Explaining discontinuity in organisational learning: a process analysis, *Organisation Studies*, Vol. 31, No. 8, pp. 1045-1068.

Berman, E. and X. Wang (2000), Performance measurement in US counties: capacity for reform, *Public Administration Review*, September/October 2000, Vol. 60, No. 5, pp. 409-420.

Bockel, J. van and M. Noordegraaf (2006), Identifying identities: performance-driven, but professional public managers, *The International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 585-597.

Bogt, H.J. ter (2004), Politicians in search of performance information? Survey research on Dutch aldermen's use of performance information, *Financial Accountability and Management*, Vol. 20, No. 3, pp. 221-252.

Bogt, H.J. ter (2008a), Recent and future management changes in local government: continuing focus on rationality and efficiency?, *Financial Accountability and Management*, Vol. 24, no. 1, pp. 31-57.

Bogt, H.J. ter (2008b), Management accounting change and new public management in local government: a reassessment of ambitions and results – an institutionalist approach to accounting change in the Dutch public sector, *Financial Accountability and Management*, Vol. 24, No. 3, pp. 209-241.

Boivard, T. and Loeffler, E. (2002), Moving from excellence models of local service delivery to benchmarking "good local governance", *International Review of Administrative Sciences*, Vol. 68, pp. 9-24.

Boivard, T. and Loeffler, E. (2003), *Public Management and Governance*, London, Routledge.

Bourne, M., Neely, A., Mills, J. and K. Platts (2003a), Implementing performance measurement systems: a literature review, *International Journal of Business Performance Management*, Vol. 5, No. 1, pp.1-24.

Bourne, M., Neely, A., Mills, J. and K. Platts (2003b),Why some performance measurement initiatives fail: lessons from the change management literature, *International Journal of Business Performance Management*, Vol. 5, No. 2-3, pp. 245-269.

Bowerman, M., Francis, G., Ball, A. and J. Fry (2002), The evolution of benchmarking in UK local authorities, *Benchmarking: an International Journal*, Vol. 9, No. 5, pp. 429-449.

Boyne, G. and J.S. Gould-Williams, Planning and Performance in Public Organisations: An Empirical Analysis, *Public Management Review*, Vol. 5, Issue 1, 2003, pp. 115-132.

Braadbaart, O. (2007), Collaborative benchmarking, transparency and performance. Evidence from the Netherlands water supply industry, *Benchmarking: An International Journal*, Vol. 14, No. 6, pp. 677-692.

Braadbaart, O. and B. Yusnandarshah (2008), Public sector benchmarking: a survey of scientific articles, 1990-2005, *International Review of Administrative Sciences*, Vol. 74, No. 3, pp. 421-433.

Brignall, T.J.S. and Ballantine, J. (2004), Strategic Enterprise Management Systems: New Directions for Research, *Management Accounting Research*, Vol. 15, pp. 225-240.

Brignall, S. and Modell, S. (2000), An institutional perspective on performance measurement and management in the 'new public sector', *Management Accounting Research*, No. 11, pp. 281-306.

Broadbent, J. and R. Laughlin (2009), Performance management systems: a conceptual model, *Management Accounting Research*, Vol. 20, pp. 283-295.

Bruijn, H. de (2002), Performance measurement in the public sector: strategies to cope with the risks of performance measurement, *The International Journal of Public Sector Management*, Vol. 15, No. 7, pp. 578-594.

Bruijn, H. de (2007), *Managing performance in the public sector*, Routledge, London and New York.

Bruijn, H. de and G.J. van Helden (2006), A plea for dialogue driven performance-based management systems: evidence from the Dutch public sector, November 2006, *Financial Accountability and Management*, Vol. 22, No. 4, pp. 405-423.

Bryman, A. (2004), *Social Research Methods*, Oxford University Press, UK.

Buchanan, D. and Storey, J. (1997), 'Role-taking and role-switching in organizational change: the four pluralities', in: McLoughlin, I. and M. Harris (eds), *Innovation, Organizational Change and Technology*, London: International Thompson Business Press.

Bunderson, J.S. and R.E. Reagans (2010), Power, Status, and Learning in Organisations, *Organisation Science*, Articles in Advance, pp. 1-13.

Burnes, B. (2000), *Managing Change: A Strategic Approach to Organizational Dynamics*, third edition, London: Pitman.

Carroll, G.R. and Hannan, M. (2001), *The demography of corporations and industries*, Princeton University Press.

Cavalluzzo, K.S. and C.D. Ittner (2004), Implementing performance measurement innovations: evidence from government, *Accounting, Organisations and Society*, Vol. 29, pp. 243-267.

CBS (2010), *Demographic Statistics per Municipality 2010*, the Hague, the Netherlands.

Central Planning Office (2011), *Measures for municipalities 2011: an analyses of the performance of local government*, pp. 143-150, The Hague.

Clegg, S., Courpasson, D. and Phillips, N. (2006), *Power and Organizations*, Thousand Oaks, CA: Sage.

Commissie Herziening Handhavingsstelsel VROM-regelgeving (July 2008), *De tijd is rijp*, The Hague, The Netherlands.

Coopey, J. (1995), The learning organisation, power, politics and ideology introduction, Management Learning, Vol. 26, No. 2, pp. 193-213.

Coopey, J. and J. Burgoyne (2000), Politics and Organisational Learning, *Journal of Management Studies*, Vol. 37, No. 6, September 2000, pp. 869-885.

Coplin, W.D., Merget, A.D. and C. Bourdeaux (2002), The professional researcher as change agent in the government performance movement, *Public Administration Review*, November/December 2002, Vol. 62, No. 6, pp. 699-711.

Crossan, M.M., Lane, H.W. and R.E. White (1999), An organisational learning framework: from intuition to institution, *The Academy of Management Review*, Vol. 24, No. 3, pp. 522-537.

Dattakumar, R. and R. Jagadeesh (2003), A review of literature on benchmarking, *Benchmarking: An International Journal*, Vol. 10. No. 3, pp. 176-209.

Dawes, S. S., Cresswell, A.M. and T.A. Pardo (2009), From "need to know" to "need to share": tangled problems, information boundaries, and the building of public sector knowledge networks, *Public Administration Review*, May/June 2009, pp. 392-402.

Dawson, P. (1994), *Organizational Change: A Processual Approach*, London: Paul Chapman.

Dooren, W. van (2006), *Performance measurement in the Flemish public sector: a supply and demand approach*, PhD thesis, Catholic University Leuven, Belgium.

Dooren, W. van, Bouckaert, G. And J. Halligan (2010), *Performance management in the public sector*, Routledge, London and New York.

Dubnick, M. (2005), Accountability and the promise of performance: in search of mechanisms, *Public Performance and Management Review*, Vol. 28, pp. 376-417.

Dutta, D.K. and M.M. Crossan (2003), Understanding change: what can we "learn" from organisational learning?, Paper presentation at the 5th International Conference on Organisational Learning and Knowledge, Canada.

Easterby-Smith, M. (1997), Disciplines of organisational learning: contributions and critiques, *Human Relations*, Vol. 50, pp. 1085-1113.

Easterby-Smith, M., Snell, R. And S. Gherardi (1998), Organisational learning: diverging communities of practice?, *Management Learning*, Vol. 29, pp. 259-272.

Easterby-Smith, M., Thorpe, R. and A. Lowe (2002), *Management Research: An Introduction*, Sage Publications.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research, *Academy of Management Review*, 14, 532-550.

Emmanuel, C., Otley, D. and K. Merchant (1990), *Accounting for Management Control* (2nd edition), T.J. Press, UK.

Fernandez, S. and H.G. Rainey (2006), Managing successful organisational change in the public sector, *Public Administration Review*, March/April 2006, pp. 168-176.

Ferreira, A. and D. Otley (2009), The design and use of performance management systems: an extended framework for analysis, *Management Accounting Research*, Vol. 20, pp. 263-282.

Folz, D.H. (2004), Service Quality and benchmarking the performance of municipal services, *Public Administration Review*, March/April 2004, Vol. 64, No. 2, pp. 209-220.

Foster, S.T. and L. Gallup (2002), On functional differences and quality understanding, *Benchmarking: An International Journal*, Vol. 9, No. 1, pp. 86-102.

Fryer, K., Antony, J. and S. Ogden (2009), Performance management in the public sector, *International Journal of Public Sector Management*, Vol. 22, No. 6, pp. 478-498.

Gold, B. (1965), Productivity Analysis and System Coherence, *Operations Research Quarterly*, Vol. 16, No. 3, pp. 287-307.

Goncharuk, A.G. and J.P. Monat (2009), A synergistic performance management model conjoining benchmarking and motivation, *Benchmarking: An International Journal*, Vol. 16, No. 6, pp. 767-784.

Grant (1996), Prospering in dynamically-competitive environments: organisational capability as knowledge integration, *Organisation Science*, Vol. 7, pp. 375-387.

Greiling, D. (2005), Performance measurement in the public sector: the German experience, *International Journal of Public Sector Management*, Vol. 22, No. 5, pp. 445-455.

Gupta, B. (2011), A comparative study of organisational strategy and culture across industry, *Benchmarking: an International Journal*, Vol. 18, No. 4, pp. 510-528.

Hatry, H. (1999), *Performance measurement: getting results*, Washington D.C. Urban Institute Press.

Haugaard, M. (2010), Power: a "family resemblance" concept, *European Journal of Cultural Studies*, Vol. 13, No. 4, pp. 419-438.

Helden, G.J. van and S. Tillema (2005), In search of a benchmarking theory for the public sector, *Financial Accountability and Management*, Vol. 21, No. 3, pp. 337-361.

Hinings, C.R. (1997), Reflections on Processual Research, *Scandinavian Journal of Management*, Vol. 13, No. 4, pp. 493-503.

Hinton, M., Francis, G. and J. Holloway (2000), Best practice benchmarking in the UK, *Benchmarking: An International Journal*, Vol. 7, No. 1, pp. 52-61.

Hodges, R. and S. Grubnic (2010), Local authority e-government partnerships in England: a case study, *Financial Accountability and Management*, Vol. 26, No. 1, pp. 42-64.

Holloway, J., Francis, G. and M. Hinton (1999), A vehicle for change? A case study of performance improvement in the "new" public sector, *The International Journal of Public Sector Management*, Vol. 12, No. 4, pp. 351-356.

Holmqvist, M. (2004), Experiential learning processes of exploitation and exploration within and between organisations: An empirical study of product development, *Organisation Science*, Vol. 15, No. 1, pp. 70-81.

Holzer, M. and K. Kloby (2005), Public performance measurement: an assessment of the state-of-the-art and models for citizen participation, *International Journal of Productivity and Performance Management*, Vol. 54, No. 7, pp. 517-532.

Holzer, M. and K. Yang (2004), Performance measurement and improvement: an assessment of the state of the art, *International Review of Administrative Science*, Vol. 70, No. 1, pp. 15-31.

Hood (1991), A public management for all seasons, *Public Administration*, Vol. 69, No. 1, pp. 3-19.

Hoogwout, M. (2007), De benchmark BWT: monitor van een sector, *Bouwregels in de Praktijk*, Vol. 2, No. 5, May 2007, pp. 16-18.

Hoogwout, M. and R. te Velde (2004), Digital building permit application: a feasibility study for a shared services solution in the Netherlands, Paper presentation at the Third International Conference, EGOV 2004, Zaragoza, Spain.

Horton, S. (2006), New public management: its impact on public servant's identity. An introduction to this symposium, *International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 533-542.

Jansen, E.P. (2008), New public management: perspectives on performance and the use of performance information, May 2008, *Financial Accountability and Management*, Vol. 24, No. 2, pp. 169-191.

Johnsen, A. (1999), Implementation mode and local government performance measurement: a Norwegian experience, *Financial Accountability and Management*, Vol. 15, pp, 41-66.

Johnsen, A. (2005), What does 25 years of experience tell us about the state of performance measurement in public policy and management?, *Public Money and Management*, January 2005, pp. 9-17.

Johnsen, A. and J. Vakkuri (2006), Is there a Nordic perspective on public sector performance measurement?, August 2006, *Financial Accountability and Management*, Vol. 22, No. 3, pp. 291-308.

Jones, R. (1999), The role of benchmarking within the cultural reform journey of an award-winning Australian municipality, *Benchmarking: An International Journal*, Vol. 6, No. 4, pp. 338-349.

Jorgensen, T.B. and B. Bozeman (2002), Public values lost? Comparing cases on contracting out from Denmark and the United States, Public Management Review, Vol. 4, pp. 63-81.

Jorgensen, T.B. and B. Bozeman (2007), Public values: an inventory, *Administration and Society*, Vol. 39, pp. 354-381.

Julnes, P. and M. Holzer (2001), Promoting the utilization of performance measures in public organisations: an empirical study of factors affecting adoption and implementation, *Public Administration Review*, November/December 2001, Vol. 61, No. 6, pp. 693-708.

Kakkuri-Knuuttila, M.-L., Lukka, K., & Kuorikoski, J. (2008), Straddling between paradigms: a naturalistic philosophical case study on interpretive research in management accounting, Accounting, *Organisations and Society*, No. 33, pp.267-291.

Kanter, R., Stein, B., and T. Jick (1992), *The Challenge of Organizational Change: How*

Companies Experience It and Leaders Guide It, New York: Free Press.

Karlöf, B. and Östblom, S. (1995), *Benchmarking: Tuottavuudella ja Laadulla Mestariksi*, Gummerus.

Kennerley, M.P., Martinez, V. and A. Neely (2004), Impact of PMS on Business Performance: A Methodological Approach, Cranfield School of Management, Bedfordshire, UK.

Kim, S. (Spring 2003), Research paradigms in organisational learning and performance: competing modes of inquiry, *Information Technology, Learning and Performance Journal*, Vol. 21, No. 1, pp. 9-18.

Kluge, A. and J. Schilling (2003), Organisational learning and learning organisations, *The Psychologist Manager-Journal*, Vol. 6, pp. 31-50.

Korsten, A.F.A. (2011), Mijn gemeente is top! Bestaan excellente overheidsorganisaties?, www.arnokorsten.nl

Kouzmin, A., Löffler, E., Klages, H. and N. Korac-Kakabadse (1999), Benchmarking and performance measurement in public sectors. Towards learning for agency effectiveness, *The International Journal of Public Sector Management*, Vol. 12, No. 2, pp. 121-144.

Kyrö, P. (2003), Revising the concept and forms of benchmarking, *Benchmarking: An International Journal*, Vol. 10, No. 3, pp. 210-225.

Kyrö, P. (2004), Benchmarking as an action research process, *Benchmarking: An International Journal*, Vol. 11, No.1, pp. 52-73.

Lapsley, I. (February 2008), The NPM agenda: back to the future, *Financial Accountability and Management*, Vol. 24, No. 1, pp. 77-96.

Lawrence, T.B., Mauws, M.K., Dyck, B. and R.F. Kleysen (2005), The politics of organisational learning: integrating power into the 4I framework, The Academy of Management Review, Vol. 30, No. 1, pp. 180-191.

Lawrence, T.B. and S. L. Robinson (2007), Ain't misbehavin: Workplace deviance as organisational resistance. Journal of Management, Vol. 33, No. 3, pp. 378-394.

Lawrence, T. B., Malhotra, N. and Morris, T. (2012), Episodic and systemic power in the transformation of professional service firms, *Journal of Management Studies*, Vol. 49, no.1, pp. 102-143.

Lee, A. (1991), Integrating positivist and interpretative approaches to organisational research, *Organisation Science*, Vol. 2, No. 4, pp. 342-365.

Leeuw, F.L. and S. van Thiel (March 2002), The Performance Paradox in the Public Sector, *Public Performance & Management Review*, Vol. 25, No. 3, pp. 267-281.

Longbottom, L. (2000), Benchmarking in the UK: an empirical study of practitioners and academics, *Benchmarking: An International Journal*, Vol. 7, No. 2, pp. 98-117.

Lukka, K. and Modell, S, (2010), Validation in interpretive management accounting research, Accounting, *Organisations and Society*, No. 35, pp. 462–477.

Magd, H. and A. Curry (2003), Benchmarking: achieving best value in public sector organisations, *Benchmarking: An International Journal*, Vol. 10, No. 3, pp. 261-286.

March, J.G. (1991), Exploration and exploitation in organisation learning, *Organisation Science*, Vol. 2, pp. 71-87.

Mausolff, C. (2004), Learning from Feedback in Performance Measurement Systems, *Public Performance and Management Review*, Vol. 28, No. 1, pp. 9-29.

McAdam, R., Hazlett, S. and C. Casey (2005), Performance management in the UK public sector. Addressing multiple stakeholder complexity, *International Journal of Public Sector Management*, Vol. 18, No. 3, pp. 256-273.

Meier, K.J., O'Toole Jr., L.J., Boyne, G.A. and R.M. Walker (2007), Strategic management and the performance of public organisations: testing venerable ideas against recent theories, *Journal of Public Administration Research and Theory*, Vol. 17, pp. 357-377.

Meier, K.J. and O'Toole Jr., L.J. (2003), Public management and educational performance: the impact of managerial networking, *Public Administration Review*, November/December 2003, Vol. 63, No. 6, pp. 689-699.

Melkers, J. and K. Willoughby (2005), Models of performance measurement use in local governments: understanding budgeting, communication, and lasting effects, *Public Administration Review*, March/April 2005, Vol. 65, No. 2, pp. 180-190.

Miles, M.B. and A.M. Huberman (1994), *Qualitative Data Analysis: An Expanded Sourcebook* (2nd edition), SAGE Publications.

Miles, R.E. and Snow, C.C. Meyer, A.D. and H.J. Coleman (1978), Organisational strategy, structure and process, *The Academy of Management Review*, Vol. 3, No. 3 (July 1978), pp. 546-562.

Mintzberg, H. and Waters, J. (1985), Of strategies, deliberant and emergent, *Strategic Management Journal*, Vol. 6, pp. 257-272.

Modell, S. (2009a), In defence of triangulation: a critical realist approach to mixed methods research in management accounting, *Management Accounting Research*, 20, pp. 208–221.

Modell, S. (2009b), Institutional research on performance measurement and management in the public sector accounting literature: a review and assessment, *Financial Accountability and Management*, Vol. 25, No. 3, pp. 277-303.

Moore, M.H. (1995), *Creating Public Value*, Cambridge, MA, Harvard University Press.

Moriarty, J.P. and C. Smallman (2009), En route to a theory of benchmarking, *Benchmarking: An International Journal*, Vol. 16, No. 4, pp. 484-503.

Moynihan, D.P. (2005), Goal based learning and the future of performance management, *Public Administration Review*, March/April 2005, Vol. 65, No. 2, pp. 203-216.

Moynihan, D.P. (2006), Managing for results in state government: evaluating a decade of reform, *Public Administration Review*, January/February 2006, pp. 77-89.

Moynihan, D.P. and P.W. Ingraham (2004), Integrative leadership in the public sector: a model of performance-information use, *Administration and Society*, Vol. 36, No. 4, pp. 427-453.

Moynihan, D.P. and S.K. Pandey (2005), Testing how management matters in an era of government by performance management, *Journal of Public Administration Research and Theory*, Vol. 15, No. 3, pp. 421-439.

Municipality of Breda (December 2007), Residential Vision Breda 2007-2010, Section Spatial Development, Breda, the Netherlands.

Municipality of Breda (December 2007), Construction Policy 2008-2011, Section Public Affairs: Building and Housing Department, Breda, The Netherlands

Municipality of Eindhoven (May 2006), Strategic Plan 2006-2009 and Annual Plan 2006, Sector Building and Housing Supervision, Eindhoven.

Municipality of Eindhoven (April 2010), Concernverslag Eindhoven 2009, Eindhoven.

Municipality of Emmen (October 2002), Woonplan Emmen 2020, Dienst Stadsontwikkeling, Afdeling Wonen.

Municipality of Groningen (2004), Handhaven met Beleid, Dienst RO/EZ, Groningen.

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, pp. 19-37.

Newell, S., Edelman, L., Scarborough, H., Swan, J. and Bresnen, M. (2003), 'Best Practice' development and transfer in the NHS: the importance of process as well as product knowledge, *Health Services Management Research*, Vol. 16, No. 1, pp. 1–12.

Newcomer, K.E. (2007), Measuring government performance, *International Journal of Public Administration*, Vol. 30, pp. 307-329.

Nicholson-Crotty, S., Theobald, N.A. and J. Nicholson-Crotty (2006), Disparate measures: public managers and performance measurement strategies, *Public Administration Review*, January/ February 2006, pp. 101-113.

Norman, R. (2002), Managing through measurement of meaning? Lessons from experience with New Zealand's public sector performance management systems, *International Review of Administrative Sciences*, Vol. 68, pp. 619-628.

Northcott, D. and S. Llewellyn (2005), Benchmarking in UK health: a gap between policy and practice?, *Benchmarking: an International Journal*, Vol. 12, No. 5, pp. 419-435.

Osborne, D. and Gaebler, T. (1992), *Reinventing government: how the entrepreneurial spirit is transforming the public sector*, Addison-Wesley

Papanioannou, T. (2007), Policy benchmarking: a tool of democracy or a tool of authoritarianism?, *Benchmarking: An International Journal*, Vol. 14, No. 4.

Parker, R. and L. Bradley (2000), Organisational culture in the public sector: evidence from six organisations, *The International Journal of Public Sector Management*, Vol. 13, No. 2, pp. 125-141.

Pettigrew, A.M. (1987), The Management of Strategic Change, Basil Blackwell, Oxford.

Pettigrew, A.M. (1990), Longitudinal field research on change: theory and practice, Organisation Science, Vol. 1, No. 3, August 1990, pp. 267-292.

Pettigrew, A.M. (1992), The character and significance of strategy process research, Strategic Management Journal, Vol. 13, Special issue: fundamental themes in strategy process research (Winter 1992), pp. 5-16.

Pettigrew, A.M. (1997), What is a Processual Analysis?, Scandinavian Journal of Management, Vol. 13, No. 4, pp. 337-348.

Pidd, M. (2005), Perversity in public service performance measurement, *International Journal of Productivity and Performance Management*, Vol. 54, No. 5/6, pp. 482-493.

Poister, H.P. and G. Streib (1999), Performance measurement in municipal government: assessing the state of the practice, *Public Administration Review*, July/August 1999, Vol. 59, No. 4, pp. 325-335.

Poister, T.H. (2003), *Measuring performance in public and nonprofit organisations*, San Francisco, Jossey-Bass.

Pollanen, R.M. (2005), Performance measurement in municipalities. Empirical evidence in Canadian context, *International Journal of Public Sector Management*, Vol. 18, No. 1, pp. 4-24.

Pollitt, C. (2005), Performance management in practice: a comparative study of executive agencies, *Journal of Public Administration Research and Theory*, June 2005, pp. 1-20.

Pollitt, C. (2006), Performance information for democracy: the missing link?, *Evaluation*, Vol. 12, No. 1, pp. 38-55.

Pollit, C. and G. Bouckaert (2004), *Public management reform: a comparative analysis*, Oxford, Oxford University Press.

Rantanen, H., Kulmala, H.I., Lönnqvist, A. And P. Kujansivu (2007), Performance measurement systems in the Finnish public sector, *International Journal of Public Sector Management*, Vol. 20, No. 5, pp. 415-433.

Rashman, L., Withers, E. and J. Hartley (2009), Organisational learning and knowledge in public service organisations: a systematic review of the literature, *International Journal of Management Reviews*, Vol. 11, No. 4, pp. 463-494.

Ridgway, V.F. (1956), Dysfunctional Consequences of Performance Measurement, *Administrative Science Quarterly*, Vol. 1, No. 2, pp. 240-247.

Rondeaux, G. (2006), Modernizing public administration: the impact on organisational identities, *International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 569-584.

Sanger, M.B. (2008), From measurement to management: breaking through the barriers to state and local performance, *Public Administration Review*, December 2008 (special issue), pp. 70-85.

Sartori, G. (1969), Politics, ideology, and belief systems, *American Political Science Review*, 63, pp. 398-411.

Schilling, J. And A. Kluge (2009), Barriers to organisational learning: an integration of theory and research, *International Journal of Management Reviews*, Vol. 11, Issue 3, pp. 337-360.

Senge, P.M. (1990), *The fifth discipline: the art and practice of the learning organisation*, New York, Doubleday Currency.

Shafritz, J. M. and A.C. Hyde (1992), *Classics of Public Administration*, Wadsworth, USA.

Sharifuddin bin Syed-Ikhsan, S.O. and F. Rowland (2004), Benchmarking knowledge management in a public organisation in Malaysia, *Benchmarking: An International Journal*, Vol. 11, No. 3, pp. 238-266.

Simon, H. A. (1946), The Proverbs of Administration, *Public Administration Review*, Vol. 17, No. 4, pp. 53-67.

Simons, R. (1995), Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal, Boston, Harvard Business School Press.

Smith, P. (1995), On the unintended consequences of publishing performance data in the public Sector, *International Journal of Public Administration*, Vol. 18, No. 2/3, pp. 277-310.

Steffy, B.D. and A.J. Grimes (1996), A critical theory of organisation science, *Academy of Management Review*, Vol. 11, No. 2, pp. 322-336.

Streib, G.D. and T.H. Poister (1999), Assessing the validity, legitimacy, and functionality of performance measurement systems in municipal governments, *The American Review of Public Administration*, Vol. 29, No. 2, pp. 107-123.

Talbot, C. (2008), Performance regimes – the institutional context of performance policies, *International Journal of Public Administration*, Vol. 31, No. 14, pp. 1569-1591.

Taylor, W. (1911), *The Principles of Scientific Management*, NY and London, Harper and brothers.

Teelken, C. (2008), The intricate implementation of performance measurement systems: exploring developments in professional-service organisations in the Dutch non-profit sector, *Review of Administrative Sciences*, Vol. 74,. No. 4, pp. 615-635.

Thiel, S. van and F.L. Leeuw (2002), The performance paradox in the public sector, *Public Performance and Management Review*, Vol. 25, No. 3, pp. 267-281.

Tillema, S. (2007), Public sector organisations' use of benchmarking information for performance improvement: theoretical analysis and explorative case studies in Dutch water boards, *Public Performance and Management Review*, Vol. 30, No. 4, pp. 496-520.

Tillema, S. and H.J. ter Bogt (2008), Dutch Local Government Auditing Committees – Effective Instrument in the Search for Improvement of Organisational Performance and Local Democracy?, paper presented at the EIASM public sector conference 2008, Amsterdam.

Triantafillou, P. (2007), Benchmarking in the public sector: a critical conceptual framework, *Public Administration*, Vol. 85, No. 3, pp. 829-846.

Vigoda-Gadot, E. and F. Yuval (2003), Managerial quality, administrative performance and trust in governance revisited. A follow up study of causality, *The International Journal of Public Sector Management*, Vol. 16, No.7, pp. 502-522.

VNG (2008), *Local government in the Netherlands*, The Hague.

Watson, G.H. (1993), *Strategic benchmarking: how to rate your company's performance against the World's best*, John Wiley & Sons, Inc., New York, NY.

Weber, M. (1922), *Economy and Society*, Edited by Max Rheinstein. Translated by Edward Shils and Max Rheinstein. New York: Simon and Schuster.

Willcocks, S.G. (2002), Adopting a multi-perspective approach to the study of public sector managerial effectiveness, *The International Journal of Public Sector Management*, Vol. 15, No. 4, pp. 262-280.

Williams, D. W. (2004), Evolution of Performance Measurement Until 1930, *Administration & Society*, Vol. 36, No. 2, pp. 131-165.

Winkelen, Van, C. and J. McKenzie (2007), Integrating individual and organisational learning initiatives: working across knowledge management and human resource management functional boundaries, *Electronic Journal of Knowledge Management*, Volume 5, Issue 4, pp. 527-538.

Wiseman, E. (2008), *The institutionalisation of organisational knowledge: learning to walk the talk*, McGill University, Montreal, Canada.

Wisniewski, M. and D. Stewart (2004), Performance measurement for stakeholders. The case of Scottish local authorities, *The International Journal of Public Sector Management*, Vol. 17, No. 3, pp. 222-223.

Wynn-Williams, K.L.H. (2005), Performance assessment and benchmarking in the public sector: an example from New Zealand, *Benchmarking: An International Journal*, Vol. 12, No. 5, pp. 482-492.

Yasin, M.M. (2002), The theory and practice of benchmarking: then and now, *Benchmarking: An International Journal*, Vol. 9, No. 3, pp. 217-243.

Yin, R.K. (2003), *Case Study Research: Design and Methods*, Sage Publications.

Zenc (October 2008), Proposal Benchmark Building and Housing Supervision Round 2009, The Hague.

Zenc (July 2009), Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): Municipality of Breda, The Hague.

Zenc (July 2009), Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): Municipality of Eindhoven, The Hague.

Zenc (July 2009), Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): Municipality of Emmen, The Hague.

Zenc (July 2009), Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): Municipality of Groningen, The Hague.

Zenc (October 2009), Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): graphs of all municipalities, The Hague.

Appendices

Appendix 1 List of articles (literature review)

1. Adcroft, A. and R. Willis (2005), The (un)intended outcome of public sector performance measurement, *International Journal of Public Sector Management*, Vol. 18, No. 5, pp. 386-400.

2. Alstete, J.W. (2008), Measurement benchmarks or "real" benchmarking? An examination of current perspectives, *Benchmarking: An International Journal*, Vol. 15, No. 2, pp. 178-186.

3. Amaral, P. and R. Sousa (2009), Barriers to internal benchmarking initiatives: an empirical investigation, *Benchmarking: An International Journal*, Vol. 16, No. 4, pp. 523-542.

4. Ammons, D.N., Coe. C. and M. Lombardo (2001), Performance comparison projects in local government: participant's perspectives, *Public Administration Review*, January/February 2001, Vol. 61, No. 1, pp. 100-110.

5. Ammons, D.N. and W.C. Rivenbark (2008), Factors influencing the use of performance data to improve municipal services: evidence from the North Carolina benchmarking project, *Public Administration Review*, March/April 2008, pp. 304-318.

6. Anand, G. and R. Kodali (2008), Benchmarking the benchmarking models, *Benchmarking: An International Journal*, Vol. 15, No. 3, pp. 257-291.

7. Andersen, B., Henriksen, B. and I. Spjelkavik (2008), Benchmarking applications in public sector principal-agent relationships, *Benchmarking: An International Journal*, Vol. 15, No. 6, pp. 723-741.

8. Andrews, R., Boyne, G.A., and R.M. Walker (2006), Strategy content and organisational performance: an empirical analysis, *Public Administration Review*, January/February 2006, pp. 52-63.

9. Askim, J., Johnsen, A. and Cristopherson K-A (2007), Factors behind organisational learning from benchmarking: experiences from Norwegian municipal benchmarking networks, *Journal of Public Administration Research and Theory*, Vol. 18, pp. 297-320.

10. Ball, A., Bowerman, M. and S. Hawksworth (2000), Benchmarking in local government under a central government agenda, *Benchmarking: An International Journal*, Vol. 7, No. 1, pp. 20-34.

11. Behn, R.D. (2003), Why measure performance? Different purposes require different measures, *Public Administration Review*, September/October 2003, Vol. 63, No. 5, pp. 586-606.

12. Berman, E. and X. Wang (2000), Performance measurement in US counties: capacity for reform, *Public Administration Review*, September/October 2000, Vol. 60, No. 5, pp. 409-420.

13. Bogt, H.J. ter (2004), Politicians in search of performance information? – survey research on Dutch aldermen's use of performance information, *Financial Accountability and Management*, Vol. 20, No. 3, pp. 221-252.

14. Bogt, H.J. ter (2008), Recent and future management changes in local government: continuing focus on rationality and efficiency?, *Financial Accountability and Management*, February 2008, Vol. 24, No. 1, pp. 31-57.

15. Bogt, H.J. ter (2008), Management accounting change and new public management in local government: a reassessment of ambitions and results – an institutionalist approach to accounting change in the Dutch public sector, *Financial Accountability and Management*, August 2008, Vol. 24, No. 3, pp. 209-241.

16. Boland, T. and A. Fowler (2000), A systems perspective of performance management in public sector organisations, *The International Journal of Public Sector Management*, Vol. 13, No. 5, pp. 417-446.

17. Bowerman, M., Francis, G., Ball A. and J. Fry (2002), The evolution of benchmarking in UK local authorities, *Benchmarking: An International Journal*, Vol. 9, No. 5, pp. 429-449.

18. Braadbaart, O. (2007), Collaborative benchmarking, transparency and performance. Evidence from the Netherlands water supply industry, *Benchmarking: An International Journal*, Vol. 14, No. 6, pp. 677-692.

19. Braadbaart, O. and B. Yusnandarshah (2008), Public sector benchmarking: a survey of scientific articles, 1990-2005, *International Review of Administrative Sciences*, Vol. 74, No. 3, pp. 421-433.

20. Brignall, S. and J. Ballantine (2004), Strategic Enterprise Management Systems: new directions for research, *Management Accounting Research*, Vol. 15, pp. 225-240.

21. Brignall, S. and S. Modell (2000), An institutional perspective on performance measurement and management in the 'new public sector', *Management Accounting Research*, Vol. 11, pp. 281-306.

22. Broadbent, J. and R. Laughlin (2009), Performance management systems: a conceptual model, *Management Accounting Research*, Vol. 20, pp. 283-295.

23. Brown, K., Waterhouse, J. and C. Flynn (2003), Change management practices. Is a hybrid model a better alternative for public sector agencies?, *The International Journal of Public Sector Management*, Vol. 16, No. 3, pp. 230-241.

24. Cavalluzzo, K.S. and C.D. Ittner (2004), Implementing performance measurement innovations: evidence from government, *Accounting, Organisations and Society*, Vol. 29, pp. 243-267.

25. Coplin, W.D., Merget, A.D. and C. Bourdeaux (2002), The professional researcher as change agent in the government performance movement, *Public Administration Review*, November/December 2002, Vol. 62, No. 6, pp. 699-711.

26. Dattakumar, R. and R. Jagadeesh (2003), A review of literature on benchmarking, *Benchmarking: An International Journal*, Vol. 10. No. 3, pp. 176-209.

27. Dawes, S. S., Cresswell, A.M. and T.A. Pardo (2009), From "need to know" to "need to share": tangled problems, information boundaries, and the building of public sector knowledge networks, *Public Administration Review*, May/June 2009, pp. 392-402.

28. De Bruijn, H. (2002), Performance measurement in the public sector: strategies to cope with the risks of performance measurement, *The International Journal of Public Sector Management*, Vol. 15, No. 7, pp. 578-594.

29. De Bruijn, H. and G.J. van Helden (2006), A plea for dialogue driven performance-based management systems: evidence from the Dutch public sector, November 2006, *Financial Accountability and Management*, Vol. 22, No. 4, pp. 405-423.

30. Fernandez, S. and H.G. Rainey (2006), Managing successful organisational change in the public sector, *Public Administration Review*, March/April 2006, pp. 168-176.

31. Ferreira, A. and D. Otley (2009), The design and use of performance management systems: an extended framework for analysis, *Management Accounting Research*, Vol. 20, pp. 263-282.

32. Folz, D.H. (2004), Service Quality and benchmarking the performance of municipal services, *Public Administration Review*, March/April 2004, Vol. 64, No. 2, pp. 209-220.

33. Foster, S.T. and L. Gallup (2002), On functional differences and quality understanding, *Benchmarking: An International Journal*, Vol. 9, No. 1, pp. 86-102.

34. Fryer, K., Antony, J. and S. Ogden (2009), Performance management in the public sector, *International Journal of Public Sector Management*, Vol. 22, No. 6, pp. 478-498.

35. Garnett, J.L., Marlowe, J. and S.K. Pandey (2008), Penetrating the performance predicament: communication as a mediator or moderator of organisational culture's impact on public organisational performance, *Public Administration Review*, March/April 2008, pp. 266-281.

36. Goncharuk, A.G. and J.P. Monat (2009), A synergistic performance management model conjoining benchmarking and motivation, *Benchmarking: An International Journal*, Vol. 16, No. 6, pp. 767-784.

37. Greiling, D. (2005), Performance measurement in the public sector: the German experience, *International Journal of Productivity and Performance Management*, Vol. 54, No. 7, pp. 551-567.

38. Grubnic, S. and M. Woods (2009), Hierarchical control and performance regimes in local government, *International Journal of Public Sector Management*, Vol. 22, No. 5, pp. 445-455.

39. Hinton, M., Francis, G. and J. Holloway (2000), Best practice benchmarking in the UK, *Benchmarking: An International Journal*, Vol. 7, No. 1, pp. 52-61.

40. Holloway, J., Francis, G. and M. Hinton (1999), A vehicle for change? A case study of performance improvement in the "new" public sector, *The International Journal of Public Sector Management*, Vol. 12, No. 4, pp. 351-356.

41. Holzer, M. and K. Kloby (2005), Public performance measurement: an assessment of the state-of-the-art and models for citizen participation, *International Journal of Productivity and Performance Management*, Vol. 54, No. 7, pp. 517-532.

42. Holzer, M. and K. Yang (2004), Performance measurement and improvement: an assessment of the state of the art, *International Review of Administrative Science*, Vol. 70, No. 1, pp. 15-31.

43. Horton, S. (2006), New public management: its impact on public servant's identity. An introduction to this symposium, *International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 533-542.

44. Jansen, E.P. (2008), New public management: perspectives on performance and the use of performance information, May 2008, *Financial Accountability and Management*, Vol. 24, No. 2, pp. 169-191.

45. Johnsen, A. (2005), What does 25 years of experience tell us about the state of performance measurement in public policy and management?, *Public Money and Management*, January 2005, pp. 9-17.

46. Johnsen, A. and J. Vakkuri (2006), Is there a Nordic perspective on public sector performance measurement?, August 2006, *Financial Accountability and Management*, Vol. 22, No. 3, pp. 291-308.

47. Jones, R. (1999), The role of benchmarking within the cultural reform journey of an award-winning Australian local authority, *Benchmarking: An International Journal*, Vol. 6, No. 4, pp. 338-349.

48. Julnes, P. and M. Holzer (2001), Promoting the utilization of performance measures in public organisations: an empirical study of factors affecting adoption and implementation, *Public Administration Review*, November/December 2001, Vol. 61, No. 6, pp. 693-708.

49. Kouzmin, A., Löffler, E., Klages, H. and N. Korac-Kakabadse (1999), Benchmarking and performance measurement in public sectors. Towards learning for agency effectiveness, *The International Journal of Public Sector Management*, Vol. 12, No. 2, pp. 121-144.

50. Kyrö, P. (2003), Revising the concept and forms of benchmarking, *Benchmarking: An International Journal*, Vol. 10, No. 3, pp. 210-225.

51. Kyrö, P. (2004), Benchmarking as an action research process, *Benchmarking: An International Journal*, Vol. 11, No.1, pp. 52-73.

52. Laise, D. (2004), Benchmarking and learning organisations: ranking methods to identify "best in class", *Benchmarking: An International Journal*, Vol. 11, No. 6, pp. 621-630.

53. Longbottom, L. (2000), Benchmarking in the UK: an empirical study of practitioners and academics, *Benchmarking: An International Journal*, Vol. 7, No. 2, pp. 98-117.

54. Magd, H. and A. Curry (2003), Benchmarking: achieving best value in public sector organisations, *Benchmarking: An International Journal*, Vol. 10, No. 3, pp. 261-286.

55. McAdam, R., Hazlett, S. and C. Casey (2005), Performance management in the UK public sector. Addressing multiple stakeholder complexity, *International Journal of Public Sector Management*, Vol. 18, No. 3, pp. 256-273.

56. McAdam, R. and L. O'Neill (2002), Evaluating best value through clustered benchmarking in UK local government: building control services, *Benchmarking: an International Journal*, Vol. 15, No. 6, pp. 438-457.

57. Meier, K.J. and L.J. O'Toole (2003), Public management and educational performance: the impact of managerial networking, *Public Administration Review*, November/December 2003, Vol. 63, No. 6, pp. 689-699.

58. Meier, K.J., O'Toole, L.J., Boyne, G.A. and R.M. Walker (2007), Strategic management and the performance of public organisations: testing venerable ideas against recent theories, *Journal of Public Administration Research and Theory*, Vol. 17, pp. 357-377.

59. Melkers, J. and K. Willoughby (2005), Models of performance measurement use in local governments: understanding budgeting, communication, and lasting effects, *Public Administration Review*, March/April 2005, Vol. 65, No. 2, pp. 180-190.

60. Micheli, P. and M. Kennerley (2005), Performance measurement frameworks in public and non-profit sectors, *Production Planning & Control*, Vol. 16, No. 2, pp. 125-134.

61. Modell, S. (2001), Peformance measurement and institutional processes: a study of managerial responses to public sector reform, *Management Accounting Research*, Vol. 12, pp. 437-464.

62. Modell, S. (2004), Performance measurement myths in the public sector: a research note, February 2004, *Financial Accountability and Management*, Vol. 20, No. 1, pp. 39-55.

63. Modell, S. (2009), Institutional research on performance measurement and management in the public sector accounting literature: a review and assessment, *Financial Accountability and Management*, Vol. 25, No. 3, pp. 277-303.

64. Moriarty, J.P. and C. Smallman (2009), En route to a theory of benchmarking, *Benchmarking: An International Journal*, Vol. 16, No. 4, pp. 484-503.

65. Moynihan, D.P. (2005), Goal based learning and the future of performance management, *Public Administration Review*, March/April 2005, Vol. 65, No. 2, pp. 203-216.

66. Moynihan, D.P. (2006), Managing for results in state government: evaluating a decade of reform, *Public Administration Review*, January/February 2006, pp. 77-89.

67. Moynihan, D.P. and P.W. Ingraham (2004), Integrative leadership in the public sector: a model of performance-information use, *Administration and Society*, Vol. 36, No. 4, pp. 427-453.

68. Mite, Y.I. (2000), Performance management model. A systems based approach to public service quality, *The International Journal of Public Sector Management*, Vol. 13, No. 1, pp. 19-37.

69. Newcomer, K.E. (2007), Measuring government performance, *International Journal of Public Administration*, Vol. 30, pp. 307-329.

70. Nicholson-Crotty, S., Theobald, N.A. and J. Nicholson-Crotty (2006), Disparate measures: public managers and performance measurement strategies, *Public Administration Review*, January/ February 2006, pp. 101-113.

71. Norman, R. (2002), Managing through measurement of meaning? Lessons from experience with New Zealand's public sector performance management systems, *International Review of Administrative Sciences*, Vol. 68, pp. 619-628.

72. Pandey, S.K. and J.L. Garnett (2006), Exploring public sector communication performance: testing a model and drawing implications, *Public Administration Review*, January/February 2006, pp. 37-51.

73. Parker, R. and L. Bradley (2000), Organisational culture in the public sector: evidence from six organisations, *The International Journal of Public Sector Management*, Vol. 13, No. 2, pp. 125-141.

74. Pidd, M. (2005), Perversity in public service performance measurement, *International Journal of Productivity and Performance Management*, Vol. 54, No. 5/6, pp. 482-493.

75. Poister, H.P. and G. Streib (1999), Performance measurement in municipal government: assessing the state of the practice, *Public Administration Review*, July/August 1999, Vol. 59, No. 4, pp. 325-335.

76. Pollanen, R.M. (2005), Performance measurement in municipalities. Empirical evidence in Canadian context, *International Journal of Public Sector Management*, Vol. 18, No. 1, pp. 4-24.

77. Pollitt, C. (2005), Performance management in practice: a comparative study of executive agencies, *Journal of Public Administration Research and Theory*, June 2005, pp. 1-20.

78. Pollitt, C. (2006), Performance information for democracy: the missing link?, *Evaluation*, Vol. 12, No. 1, pp. 38-55.

79. Rantanen, H., Kulmala, H.I., Lönnqvist, A. and P. Kujansivu (2007), Performance measurement systems in the Finnish public sector, *International Journal of Public Sector Management*, Vol. 20, No. 5, pp. 415-433.

80. Rondeaux, G. (2006), Modernizing public administration: the impact on organisational identities, *International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 569-584.

81. Rondo-Brovetto, P. and I. Saliterer (2007), Comparing regions, cities and communities: local government benchmarking as an instrument for improving performance and competitiveness, *The Innovation Journal: the Public Sector Innovation Journal*, Vol. 12, No. 3, Article 13, pp. 1-18.

82. Sanger, M.B. (2008), From measurement to management: breaking through the barriers to state and local performance, *Public Administration Review*, December 2008 (special issue), pp. 70-85.

83. Sharifuddin bin Syed-Ikhsan, S.O. and F. Rowland (2004), Benchmarking knowledge management in a public organisation in Malaysia, *Benchmarking: An International Journal*, Vol. 11, No. 3, pp. 238-266.

84. Streib, G.D. and T.H. Poister (1999), Assessing the validity, legitimacy, and functionality of performance measurement systems in municipal governments, *The American Review of Public Administration*, Vol. 29, No. 2, pp. 107-123.

85. Talbot, C. (2008), Performance regimes – the institutional context of performance policies, *International Journal of Public Administration*, Vol. 31, No. 14, pp. 1569-1591.

86. Teelken, C. (2008), The intricate implementation of performance measurement systems: exploring developments in professional-service organisations in the Dutch non-profit sector, *Review of Administrative Sciences*, Vol. 74,. No. 4, pp. 615-635.

87. Tillema, S. (2007), Public sector organisations' use of benchmarking information for performance improvement: theoretical analysis and explorative case studies in Dutch water boards, *Public Performance and Management Review*, Vol. 30, No. 4, pp. 496-520.

88. Triantafillou, P. (2007), Benchmarking in the public sector: a critical conceptual framework, *Public Administration*, Vol. 85, No. 3, pp. 829-846.

89. Van Bockel, J. and M. Noordegraaf (2006), Identifying identities: performance-driven, but professional public managers, *The International Journal of Public Sector Management*, Vol. 19, No. 6, pp. 585-597.

90. Van Dooren, W. (2005), What makes organisations measure? Hypotheses on the causes and conditions for performance measurement, *Financial Accountability and Management*, Vol. 21, No. 3, pp. 363-383.

91. Van Helden, G.J. and S. Tillema (2005), In search of a benchmarking theory for the public sector, *Financial Accountability and Management*, Vol. 21, No. 3, pp. 337-361.

92. Van Helden, G.J., Johnsen, A. and J. Vakkuri (2008), Distinctive research patterns on public performance measurement of public administration and accounting disciplines, *Public Management Review*, Vol. 10, Issue 5, pp. 641-651.

93. Van Thiel, S. and F.L. Leeuw (2002), The performance paradox in the public sector, *Public Performance and Management Review*, Vol. 25, No. 3, pp. 267-281.

94. Vigoda-Gadot, E. and F. Yuval (2003), Managerial quality, administrative performance and trust in governance revisited. A follow up study of causality, *The International Journal of Public Sector Management*, Vol. 16, No.7, pp. 502-522.

95. Wisniewski, M. and D. Stewart (2004), Performance measurement for stakeholders. The case of Scottish local authorities, *The International Journal of Public Sector Management*, Vol. 17, No. 3, pp. 222-223.

96. Willcocks, S.G. (2002), Adopting a multi-perspective approach to the study of public sector managerial effectiveness, *The International Journal of Public Sector Management*, Vol. 15, No. 4, pp. 262-280.

97. Williams, D.W. (2004), Evolution of performance measurement until 1930, *Administration and Society*, Vol. 36, No. 2, pp. 131-165.

98. Wynn-Williams, K.L.H. (2005), Performance assessment and benchmarking in the public sector: an example from New Zealand, *Benchmarking: An International Journal*, Vol. 12, No. 5, pp. 482-492.

99. Yang, K. and J.Y. Hsieh (2007), Managerial effectiveness of government performance measurement: testing a middle-range model, *Public Administration Review*, September/October 2007, pp. 861-879.

100. Yasin, M.M. (2002), The theory and practice of benchmarking: then and now, *Benchmarking: An International Journal*, Vol. 9, No. 3, pp. 217-243.

Appendix 2 Analysis of the studies: research questions and conclusions

Author	Date	Context (where) (why and when)	Content (what)	Process (how)	Conclusions
Adcroft and Willis	2005			Do performance measurement systems have an influence on improving services?	Current systems of performance measurement in the public sector are unlikely to have a significant influence on improving services. The most likely outcomes of these systems are further commodification (the transformation of relationships into quasi-commercial relationships) of services and de-professionalisation of public sector workers. The authors indicate to six systemic problems of using performance measurement for service improvement (see page 394-396).
Alstete	2008		What are the current perceptions of company employees regarding the use and understanding of the terms "benchmark" as in performance measurement and "benchmarking" as in measurement followed by identification of best practices for improvement.		There is a misunderstanding between the commonly used terms relating to performance measurement benchmarks and true benchmarking where identification of process leaders is only the first step in a complete process that then identifies best practices to be adapted. This paper reveals that a more precise terminological use of true benchmarking practices should be promoted and used by management leaders, educators and writers.
Amaral and Sousa	2009	What are typical barriers that occur with internal benchmarking in a manufacturing plant?			Barriers to internal benchmarking initiatives of different types are found: organisational barriers (people, culture, and context), benchmarking project management barriers (planning and implementation, leadership, and business

					pressures) and benchmarking data barriers (difficulty to access/compare data).
Ammons, Coe and Lombardo	2001		How do officials of participating local governments assess the value of their involvement in performance comparison projects, including related costs and benefits?		Benchmarking offers a tool, not a promise. The organisers of comparative projects must drive this point home or risk unrealistic expectations. Participating units that expect obvious and easy solutions for inefficiencies and service shortcomings face inevitable frustration. Successful bench markers make plans for the *use* of their new tool, and they carry those plans out. Their less successful counterparts often incorrectly assume that good results will accrue simply from *having* the tool.
Ammons and Rivenbark	2008			What are patterns of performance measurement use in North Carolina (US) cities? If cities participating in the project make greater use of performance measures, why is this so? And why do some of the participants in the project use the data to influence operations more than other participants do?	The experience of 15 participating municipalities suggests that the likelihood that performance data will be used to influence operations is enhanced by the (1) collection of and reliance on higher-order measures, especially efficiency measures, rather than simply workload or output measures; (2) the willingness of officials to embrace comparison with other governments or service producers; and (3) the incorporation of performance measures into key management systems.
Anand and Kodali	2008		What is the fundamental classification scheme of benchmarking within companies and thereby the unique benchmarking models that are developed for each type of		The benchmarking analysis of various benchmarking models revealed that each models differs in terms of number of factors – number of steps involved, number of phases, type of benchmarking it is applied, etc. The best practices identified from this process have been categorised into different phases and the proposed model

			benchmarking?		consists of 12 phases which includes about 54 steps (both common as well as best practices in benchmarking) identified during the process. An analysis of the taxonomy of benchmarking models revealed that benchmarking as a tool has more practical or industrial utility than academic/research utility.
Andersen, Henriksen, and Spjelkavik	2008	What is the range of benchmarking applications that can be used in a principal-agent relationship setting often found in the public sector?			The voluntary, improvement-oriented type of benchmarking has the highest potential for generating benefits to the average public sector organisation. The paper demonstrates that also the compulsory type of benchmarking can be useful. The principal-agent theory suitably describes the context within which compulsory benchmarking can be put to useful use.
Andrews, Boyne and Walker	2006		Is strategy content a key determinant of organisational performance in the public sector?		According to the authors strategy content comprises two dimensions: strategic stance (the extent to which an organisation is a prospector, defender, or reactor) and strategic actions (the relative emphasis on changes in markets, services, revenues, external relationships, and internal characteristics). The statistical results show that strategy content matters. Organisational performance is positively associated with a prospector stance and negatively with a reactor stance. Furthermore, local authorities that seek new markets for their services are more likely to perform well. These results suggest that measures of strategy content must be included in valid theoretical and empirical models of organisational performance in the public sector.
Askim,	2007	To what extent is			Municipalities do obtain organisational learning

Johnsen and Cristopherson		organisational learning from benchmarking conditioned by the composition of benchmarking networks, internal organisational processes, political factors, and history?			from benchmarking but care must be taken when organisational learning is conceptualised and assessed. Learning should incorporate aiding agenda setting and decision making, as well as changes. Factors such as network and administrative characteristics and management and political participation are found to influence learning outcomes. There are also indications that learning from benchmarking is subject to politics. Notably, non-socialist political regimes are less receptive to organisational learning than other regimes and political competition enhances organisational learning from benchmarking.
Ball, Bowerman, Hawksworth	2000		What is the role of performance monitoring in local government?		What is clearly missing from the existing interpretation of benchmarking in the public sector is a recognition of the impact of the policy process. The corollary is a conflation of two distinct views of benchmarking: benchmarking as a rigorous and challenging scrutiny of local government processes; and benchmarking as an instrument of central government control.
Behn	2003		Which performance measures should public managers deploy to meet the managerial purposes to which performance measures might contribute?		As part of their overall management strategy, public managers can use performance measures to evaluate, control, budget, motivate, promote, celebrate, learn, and improve. Unfortunately, no single performance measure is appropriate for all eight purposes. Consequently, public managers should not seek the one magic performance measure. Instead, they need to think seriously about the managerial purposes to which performance measurement might contribute and how they might deploy these measures. Only then can they select measures with the characteristics necessary to help achieve each purpose. Without

					at least a tentative theory about how performance measures can be employed to foster improvement (which is the core purpose behind the other seven), public managers will be unable to decide what should be measured.
Berman and Wang	2000	To what extent do counties implement performance measurement? Which capacities must be present for different levels of implementation and success? What can counties do to increase their capacity for performance measurement? And, what is the effect of county structure and functions on the use of performance measurement?			This study finds that about one-third of counties use performance measurement and that about one-fifth of these have a high level of use. Among those that use performance measurement, about one-third (31 percent-39 per cent) have an adequate level of capacity, although the level of capacity decreases sharply among low users (9 per cent). Capacity requires that jurisdictions are able (1) to relate outputs to operations; (2) to collect timely data; have (3) staff capable of analysing performance data; (4) adequate information technology; and support from (5) department heads and (6) elected officials. This study finds that the success of performance measurement is greatly affected by counties' underlying organisational capacities.
Bogt, ter	2004			Is the aldermen's opinion on and use of various sources of performance information related to the policy fields for which they are responsible?	Many aldermen see little value in the output-oriented performance information that is available in the planning and control documents of their organisations and they use it only infrequently. In general, all aldermen seemed to prefer rich, verbal information to sources of written information, probably because they work in a relatively complex and uncertain political environment. The relationship between aldermen's use of performance information and the main policy fields in their portfolios is not straightforward. The survey shows that aldermen obtain much

					information on performances from other sources besides budgets, annual reports and interim reports. It is unclear whether technical improvements to planning and control instruments will satisfy those local politicians who are especially interested in 'political efficiency' or who think that many aspects of government activities, outputs, and outcomes cannot easily be expressed in quantitative terms. Moreover, as the position of aldermen is, in some respects, similar to that of top managers in for-profit organisations, aldermen might prefer qualitative, rich information to historical, standard, written, and numeric output information in planning and control documents.
Bogt, ter	2008 (February)	What has been the effect of introducing several NPM related management changes on the functioning and effectiveness of Dutch local government?			The empirical research clearly shows that the organisations researched have introduced many management changes in recent years. In addition, the findings indicate that Dutch local government will introduce further change initiatives. The above discussion suggested several changes that organisations might introduce in the near future (for example benchmarking, HRM, transparency and PPP). For the near future, there seem to be indications that the reforms likely to be introduced will emphatically be intended to increase performance, efficiency, effectiveness and economic rationality in general, or at least gives that impression. This emphasis may be due to budgetary problems, but it could also be part of socially and politically rational behaviour on the part of politicians and professional managers. All this seems to suggest that in the future, too, politicians and managers will pay attention to efficiency and effectiveness, the traditional focus of PMI/NPM, although quite a number of

					politicians and managers are critical of the various NPM-type changes to Dutch local government in recent years.
Bogt, ter	2008 (August)	Which changes were made to management accounting systems since 1985 in Dutch local Government and what have been the actual effects in government organisations?			Many local government organisations introduced output and outcome-oriented planning and control, customer orientedness, competence management, quality models and multidimensional performance management, or strengthened the role of such instruments since 1985. This paper indicates that an institutionalist perspective enables one to study change processes in organisations and to observe issues and developments that might not be noticed when a more functional and short-term perspective is chosen. It seems that the accounting changes, even though they were not a success in a technical sense, did bring about some effects in organisational culture and individual behaviour that are in keeping with the ideas of NPM, i.e. a greater focus on performance, external stakeholders, and a business-like attitude. Accounting changes can be regarded as a change in the rules in the organisations involved. The research seems to indicate that the routines in the organisations have also changed, but at a slower pace than the rules, and perhaps in unexpected ways. The findings suggest that social factors and structures influence the accounting change process in the organisations to a considerable extent.
Boland and Fowler	2000			What is the potential role of influence diagrams and dynamic simulation models (within the context	Due consideration of systemic relationships, within the context of the control location/action matrix model, is firmly endorsed as potentially providing a framework within which attitudes

				of performance management and system dynamics) as a potential means of unravelling the complex behaviour which can often arise in the presence of interactive cause-effect loops?	and policies towards performance management in the public sector can be reconsidered and possibly refocused to reflect outcome attainment clearly. It is accepted that thinking through the full run of cause and effect sequences, arising in the interconnected dynamic feedback loops which are encountered in the domain of public sector management, may be a daunting task in practice. However, the process of qualitative mapping and modelling using influence diagrams, and quantitative modelling leading to dynamic simulation, can offer considerable assistance in this respect. These tools can subsequently reveal, at the policy design and reformulation stage, the likely behaviour of the system, following implementation of performance management initiatives or the experience ``shocks to the system'' arising from an uncontrollable environment.
Bowerman, Francis, Ball and Fry	2002			What are the reasons for benchmarking in the public sector in UK local authorities?	The reasons for benchmarking in the public sector are confused; pressures for accountability in the public sector may militate against real performance improvement; and an appropriate balance between the use of benchmarking for control and improvement purposes is yet to be achieved.
Braadbaart	2007	Can collaborative benchmarking boost the performance of public sector organisations in the Netherlands water supply industry?			Benchmarking immediately enhanced transparency, but only affected utility economic performance after benchmarking information entered the public domain. This confirms that benchmarking enhances transparency and performance. The findings do not support the yardstick regulation hypothesis that utility managers will only tighten financial discipline

					when benchmarking is embedded in a regime of managed competition.
Braadbaart and Yusnandarshah	2008		What has been the evolution of the academic literature on Public Sector Benchmarking (PSB) spanning the period 1990–2005?		Between 1990 and 2005 PSB evolved into a mature and strongly international field of research. A theoretical and conceptual rift runs through the literature, with those advocating PSB as a tool for managed competition on one side, and those promoting benchmarking as a voluntary and collaborative learning process on the other. A first challenge facing future PSB researchers is that of closing the gap between the managed and voluntary benchmarking perspectives; a second challenge concerns empirical tests that capture the effects of different benchmarking regimes on the performance of public sector providers.
Brignall and Ballantine	2004		What is the relationship between performance measurement and management (PMM) and Strategic Enterprise Management (SEM)?		SEM's are at best a partial solution to the quest for performance improvement, a solution which implies more questions about the context, content and process of organisational change and performance improvement. Future research on what determines the success of SEM initiatives might usefully combine two research methods. Large sample multivariate statistical studies could be used to identify the "what" of the complementarities among their component parts while an associated set of longitudinal case studies might help answer the questions about the process and context of successful change and performance improvement.
Brignall and Modell	2000	What are the implications of institutional theory for the successful implementation of			The authors show that the differing nature of the interrelationships between these three key stakeholders will influence the extent to which performance

		multidimensional performance measurement and management in the public sector?			measurement in the focal service-provider organisations will be *balanced* and *integrated*. They also discuss the influence of these core concepts on the possibilities of achieving some balance between the stakeholder interests examined in the overall control of provider organisations. The authors argue that studies of managerial choice constitute a useful starting point for analysing how PM practices change in highly institutionalised settings, such as the public sector.
Broadbent and Laughlin	2009	What are the underlying factors that influence the nature of any PMS?			The analysis leads to the development of a 'middle range' model of the alternative nature of PMS lying on a continuum from 'transactional' at one end to 'relational' at the other built on respectively underlying instrumental and communicative rationalities and guided by a range of contextual factors (building on the work and insights of primarily Otley (1999) and Ferreira and Otley (2005, 2009).
Brown, Waterhouse and Flynn	2003		Delivers a hybrid model of "new public management" more favourable outcomes than a model focused on cost reduction in change management processes?		The research study demonstrates that a hybrid model of NPM has been able to deliver superior outcomes for employees as well as political and public stakeholders. Such a model has taken advantage of the best of both bureaucracy and more egalitarian styles of management through the development of a relational culture while balancing its outcomes through the adoption of a balanced scorecard that focuses on finance, operational efficiency, customer/stakeholder satisfaction and human resources management.
Cavalluzzo and Ittner	2004	Which factors influence the development, use, and			The authors find that organisational factors such as top management commitment to the use of

		perceived benefits of results-oriented performance measures in government activities?			performance information, decision-making authority, and training in performance measurement techniques have a significant positive influence on measurement system development and use. They also find that technical issues, such as information system problems and difficulties selecting and interpreting appropriate performance metrics in hard-to-measure activities, play an important role in system implementation and use. The extent of performance measurement and accountability are positively associated with greater use of performance information for various purposes. However, the authors find relatively little evidence that the perceived benefits from recent mandated performance measurement initiatives in the US government increase with greater measurement and accountability. Finally, they provide exploratory evidence that some of the technical and organisational factors interact to influence measurement system implementation and outcomes, often in a complex manner.
Coplin, Merget and Bourdeaux	2002			What should be the role of professional researchers in the government performance movement?	Although academic and professional publications give the impression that performance measurement is a growing government practice, in actuality the use of this technology is not as deep as widespread as if may appear. Even when performance measures are used, governments rarely integrate them into planning, budget, personnel, and other management processes. Most professional researchers located primarily in academic institutions, but also in research and government organisations, approach performance measurement as though governmental officials, elected or otherwise, are already sold on its usefulness. Instead, they need to function as

					"change agents," using a variety of strategies to gain acceptance and understanding of the strengths and limitations of performance measurement.
Dattakumar and Jagadeesh	2003		Which papers are available on benchmarking (1980 till June 2002) and which schemes of classification have been used?		Considering the gamut of publications it can be said that the benchmarking technique has seen a steady growth and appears to be heading towards maturity level. The present review of literature on benchmarking, carried out as a part of on-going research, has identified certain issues which have not been satisfactorily addressed or not been addressed at all. These are: (1) cost aspects of benchmarking; (2) duration of benchmarking exercise; (3) human resources in benchmarking activities and; (4) selecting benchmarking partner.
Dawes, Cresswell and Pardo	2009		What are the challenges that Public Sector Knowledge Networks (PSKN) face?		1. While the problems of starting and sustaining PSKNs are formidable, they are not beyond the capabilities of astute, strategic, and tactically adept network builders; 2. IT considerations must be appreciated as nested within a variety of organisational, sociological, ideological, and political contexts that all need considerable attention; 3. Political leaders and public managers need to invest in developing as fundamental public management skills a broad and deep understanding of and capability for engaging with the *Realpolitik* of sharing knowledge and information in networks. Network development processes that emphasise early, open dialogue and examination of

					assumptions and expectations do better than those that rush forward with a fixed IT solution in mind. Those that adapt and learn from experience are more likely to succeed in achieving their substantive project and networking goals. Finally, to be sustainable as organisational forms, knowledge networks need some legal foundation, access to resources, supportive policies, and innovative forms of leadership.
De Bruijn	2002			What do we know about the perverse effects of performance measurement? How can performance measurement be shaped so as to minimise its perverse effects?	Perverse effects of performance measurement: • Prompts game playing • Adds to internal bureaucracy • Blocks innovations • Blocks ambitions • Performance measurement professionalism • Kills system responsibility • Punishes good performance The author presents five successive strategies aimed at preventing perverse effects where possible: 1. Tolerating competing product definitions; 2. Banning a monopoly on interpreting production figures; 3. Limiting the functions of and forums for performance measurement; 4. Strategically limiting the products that can be subjected to performance measurement; and 5. Using a process perspective of performance measurement in addition to a product perspective.
De Bruijn and	2006			What factors explain the	Performance measurement systems in the public

Van Helden				success or lack of success of performance measurement systems in the public sector?	sector will be successful if these systems are developed and used in an interactive way between managers and professionals. Without a behavioural-oriented approach, systems of performance management can create strong incentives for perverse behaviour and might therefore be a victim of the Law of Decreasing Effectiveness: more control, appraisal and sanctions will lead to more perverse effects and reduced effectiveness.
Fernandez and Rainey	2006			To which factors should change managers and change participants pay attention in order to successfully implement a change process?	The factors and propositions offered in this article should serve not as a road map but as a compass for practitioners seeking to find their way amid the sustained, persistent, and challenging pressures for change they confront daily. Factors and propositions suggested: (1) ensure the need; (2) provide a plan; (3) build internal support for change and overcome resistance; (4) ensure top management support and commitment; (5) build external support; (6) provide resources; (7) institutionalise change; and (8) pursue comprehensive change.
Ferreira and Otley	2009		Which questions to include in the performance management systems framework to provide an overview of the major performance management issues within an organisation?		The authors put forward the performance management systems framework as a research tool for describing the structure and operation of performance management systems (PMSs) in a more holistic manner. In particular, it elaborates the 5 questions of Otley's (1999) performance management framework into 12 questions. Anecdotal evidence suggests that the extended framework provides a useful research tool for those wishing to study the design and operation of performance management systems by providing a template to help describe the key

					aspects of such systems.
Folz	2004	How can local officials select benchmarking partners whose best practices have the most potential for applicability and success in improving service performance?			This research suggests that public administration scholars can advance the benchmarking process - particularly, the search for best practice partners - by devising frameworks that enable local officials to distinguish the different levels of quality in municipal service provision. The study shows that a quality-of-service framework for municipal services can advance local decision making about what citizens and stakeholders expect and will support in terms of input service quality. It also can help local officials identify benchmarking partners that provide a service at the desired level of quality.
Foster, Gallup	2002	How is quality improvement perceived in western state companies that manufacture a variety of products?			Differences exist in how people with different functional job classifications view quality and quality improvement efforts. Functional perspective (engineering, operations, strategic management, marketing, financial and human resources) is an important reason for differing perceptions. Communication problems exist between people in the different functions. This problem is expected to be especially pronounced in companies that are functionally organised. To overcome this problem, companies need to involve the entire staff in improving the company. In this context, cross-functional teams can take advantage of these multiple perspectives to improve quality.
Fryer, Antony and Ogden	2009		What is the state of performance management in the public sector?		The expected improvements in performance, accountability, transparency, quality of service and value for money have not yet materialised in

					the public sector. There are three classes of problems with performance management in the public sector: 1. Technical problems (relate to the indicators and the data, their collection, interpretation and analysis); 2. Systems problems (integrating performance systems with the existing systems, a lack of strategic focus which encourages short-termism, the ambiguity of performance objectives, sub-optimization and the cost of performance management); 3. Involvement problems (people issues and their involvement in the performance management system). Externally imposed restructurings and reorganisations restrict the successful implementation of performance management.
Garnett, Marlowe and Pandey	2008	What is the role that communication plays in achieving organisational performance?			The analysis supports the claim that communication acts as a meta-mechanism for shaping and imparting culture in mission-oriented organisational cultures, thereby influencing performance. In particular, task orientation, feedback, and upward communication have positive effects on perceived organisational performance in mission-oriented organisations but potentially negative effects on performance in rule-oriented cultures.
Goncharuk and Monat	2009		What are the strengths and weaknesses of conventional benchmarking approaches to productivity		It is found that most benchmarking efforts are hampered by resistance of employees to change. It is therefore concluded that benchmarking efforts could be enhanced by integrating employee motivation/behaviour programs with the

			maximization in the private sector?		benchmarking efforts. The conjoining of internal benchmarking, external benchmarking, and employee motivation/behaviour programs should substantially enhance the results of productivity improvement programs based upon benchmarking.
Greiling	2005			How has performance measurement within the German public sector been used?	The German public sector can be described as a late starter with respect to performance measurement. The full potential performance measurement may offer in the opinion of its supporters is not realised. The front runners of performance measurement are local governments. Voluntary inter-administrative comparison circles are the most frequently used instrument, followed – to a much lesser extent – by quality-management initiatives and performance-indicator-based contracting. The experience with and the acceptance of comparison circles have been mixed. According to the author further research is needed into the transaction and opportunity costs of performance measurement and into the conditions under which the performance measurement can support an organisational learning process.
Grubnic and Woods	2009			What is the extent of control exhibited by central government over local government through the best value (BV) and comprehensive performance assessment (CPA) performance regimes (using the	Comprehensive performance assessment (CPA) represents a more sophisticated performance regime than best value (BV) in the governance of local authorities by central government. In comparison to BV, CPA involved less scope for dialogue with local government prior to introduction, closer inspection of and direction of support toward poorer performing authorities, and more alignment to government priorities in the weightings attached to service blocks.

				dimensions of hierarchical control as articulated by Ouchi)?	
Hinton, Francis, Holloway	2000		What are the disincentives to benchmarking activity experienced by practising bench markers, as well as the factors which inhibit the initial take-up of this technique?		A great deal of benchmarking activity can be described as ``results'' benchmarking as opposed to ``process'' benchmarking. While it is harder to develop process measures they can prove far more valuable in improving performance and help to overcome problems such as comparability. It is strongly recommended that: a. Benchmarking should not be restricted merely to comparisons of results but include an examination of the underlying causal processes; b. Bench markers should pay at least as much attention to the organisational climate as to the technical or formal steps taken; c. When benchmarking activities are planned, attention should be paid to training in team working, communications and change management equal to technical skills associated with the steps of benchmarking; d. Prior experience seems to be an important catalyst, so novice bench markers could consider co-opting someone with experience, working with an experienced partner or joining a benchmark club or network.
Holloway, Francis and Hinton	1999	Can a single approach to performance improvement be responsible for significant organisational transformation?			The authors argue that complex approaches to performance improvement such as benchmarking, however technically powerful they may be, are only as effective as the people who apply them and their compatibility with the organisational context in which they are used. The contribution

					of such methods is often difficult to separate from other variables. In addition to internal organisational characteristics, external contextual factors play an important part both in establishing a need to use such approaches, and encouraging commitment to their use.
Holzer and Kloby	2005			How are citizens adding meaning to the performance measurement process? What is the state-of-the-art of public performance measurement?	While there are challenges associated with implementing systems of performance measurement, evidence shows that including citizens in the process adds value to the overall process. The authors highlight that citizen-driven government requires strong systems for public performance measurement.
Holzer and Yang	2004			What is the experience with citizen-driven government performance measurement?	Exemplary practices are evidence that citizen-driven government performance measurement is conceptually sound. The article concludes that the essential question in the future is how government can move to full adoption and implementation of citizen-driven, data-driven decision-making.
Horton	2006	What is the impact of new public management on public servant's identity?			NPM is impacting on public officials as their roles and the work they do, the ways in which they are managed, their relationships with the public and the criteria by which they are assessed, both internally and externally, are continually evolving. As the public service is losing its specificity and its unique role and mode of operation, being a civil servant or public official no longer has such a distinct identity. Individual civil servants are adjusting their perceptions of the collective identity, the public's perception of that identity and their own self-identity. What appears to be the case is that in spite of all the changes much of

					the traditional public administration cultures remain. The article also highlights the contribution that cultural and social theories, drawn from anthropology and organisational psychology, make to an understanding of the processes by which public servants' identity are formed and changed.
Jansen	2008			How do the perspectives on performance of politicians and managers in the public sector affect the extent to, and the way in which, they use the performance information that is now available as a result of the decision to adopt NPM (The Netherlands)?	Politicians have a different perspective on performance, as compared to the internal process perspective and output perspective which are implied in NPM. Managers, especially Production Managers, have an internal perspective, whereas politicians seem to have a citizen perspective and a financial perspective on performance. The citizen perspective is typical of governmental organisations and it can limit the implementation of more business-like ways of running such organisations. For politicians, there needs to be an incentive to use information about internal processes and outputs.
Johnsen	2005		What does 25 years of experience tell us about the state of performance measurement in public policy and management?		Performance indicators (PIs) have diverse functions for different stakeholders over the life-cycle of a public policy, and the search for better PIs is an on-going effort. However, instead of seeing the running down, proliferation and strategic use of performance information as dysfunctional, these effects are probably the unavoidable outcomes of functional and effective performance measurement systems in open societies and competitive democracies. PIs may effectively create 'creative destruction' of the present political or managerial status quo. Thus,

					PIs in political competition may be as important as prices in market competition.
Johnsen and Vakkuri	2006	Is there a Nordic perspective on public sector performance measurement and what are potential implications of such a perspective for performance measurement in public management?			The Nordic perspective has emphasised a stakeholder approach, bottom-up processes, loose couplings (albeit not necessarily decoupling), and measurements for organisational learning rather than individual financial rewards. This could mean that in the Nordic countries performance measurement may be used relatively more for dialogue and learning than for management control. Because the Nordic model depends on economic flexibility, social innovation, and political compromises the Nordic perspective may favour a homeostatic model more than a cybernetic model. The Nordic countries as such may also include for instance the Netherlands as a politically and culturally similar country.
Jones	1999			What is the use of benchmarking as part of a quality oriented cultural change program?	Comparative benchmarking provides the trigger (by identifying areas of poor performance), but process knowledge provides the means of identifying how performance can be improved.
Julnes and Holzer	2001			Which factors affect the utilisation of performance measurement in public organisations?	Policy adaption is driven more heavily by factors from rational and technocratic theory, whereas actual implementation is influenced by factors addressed by political and cultural considerations. The findings suggest that utilisation can be effected by doing the following: a. Conduct an assessment of organisation's "readiness" to develop and implement performance measures; b. This may reveal the level of knowledge in the organisation about the usefulness of performance-measurement information, the

					level of support for performance measurement, and the condition of the organisation as it relates to its culture, resources, and expertise; c. Identify and involve the organisation's internal and external interest groups; d. Involve employee unions; e. Support the adoption of performance measures even if the organisation is not able to implement performance measures in a short period of time. The awareness and culture that the adoption of performance measures can create may help improve the chances for implementation later on; f. Emphasise the need to develop a 'performance improvement" culture.
Kouzmin, Löffler, Klages and Korac-Kakabadse	1999	What are the vulnerable points of benchmarking?			Technical problems, scepticism about usefulness and the appropriateness of transferring putative private sector competencies into public administration and the resistance in accepting organisational change as a necessary consequence of benchmarking exercises in the public sector, prevent the widespread acceptance and use of benchmarking in public sectors. Nevertheless, there are some encouraging examples of benchmarking within the public sector. A preliminary attempt at drawing some general conclusions about expanding German and other European benchmarking experiences involves nine emergent propositions (see page 131-134).
Kyrö	2003		How do the most recent developments in the content, forms and targets		It is argued that the need for re-conceptualising is due both to the appearance of three new forms of benchmarking (i.e. a competence benchmarking, a

			of benchmarking revise its concept and classifications?		global benchmarking and, as an option, a networking benchmark) and new fields of benchmarking (i.e. public- and semi-public sectors, as well as small firms). Public-sector organisations and small businesses, in particular, might benefit from the option of networking benchmarking. The latest research in education might provide some new insights into learning in networks.
Kyrö	2004			Is it possible to adopt an action research approach in order to advance understanding of the benchmarking process as an interplay between scientific and practical knowledge?	The results indicate that benchmarking might be regarded as a special kind of action research; however, in that case, more attention would need to be addressed to the preliminary planning, observation, reflection and the use of theoretical frames, all of which are essential for distinguishing a scientifically-conducted action research process from practical work. Thus, adopting an action research approach might improve the implementation as well as the theoretical foundation of the benchmarking process.
Laise	2004		What are the advantages, in terms of greater flexibility and realism, connected to the application of the multi-criteria methodology founded on the notion of outranking?		Outranking methods make it possible to deal with multi-criteria benchmarking and avoid the shortcomings of the traditional methods based on the average aggregate mono-criterion. If applied to the measurement of learning capability, they are a complete alternative to the traditional approach. They can support the behavioural theory of organisational analysis. This methodology solves the multi-criteria benchmarking problem without using the averaging rule adopted by the traditional benchmarking approach.

Longbottom	2000	What is the impact of benchmarking projects on performance improvement in the UK, and what are the critical factors for transfer of best practices between organisations?			a. The selection of projects is rarely based on or flows from the strategic planning process. This leads to sub-optimal results and in some instances misuse of resources b. The focus of projects is very narrow, and customer processes are being neglected c. Benchmarking projects and performance improvement are positively related d. Support is found for improvement in organisation culture through staff development and learning e. Critical factors for transfer of best practices between organisations: best projects are established from the strategic planning process, they focus on establishing best practices through understanding process differences, they have particular team characteristics evident, and there exists an established organisation pre-culture which supports such initiatives
Magd and Curry	2003	Can best value be achieved in public-sector organisations through the implementation of benchmarking? Is benchmarking useful in public sector organisations?			In order for benchmarking to be successful in public-sector organisations, it is important to have a full commitment to continuous improvement, an ability to learn from others, and a commitment to implement improvement. Furthermore, organisations should have a supportive management team, ready access to partners who have solved the problem and a knowledgeable benchmarking team capable of solving any problems identified.
McAdam, Hazlett and Casey	2005	How can the multiple, and sometimes conflicting, stakeholder requirements, be			It is suggested that the initial starting point and key driver for performance measurement should rather be focused on “stakeholder windows” (Wisniewski and Stewart, 2004). In this approach

		represented and satisfied within a performance management approach in a large government department? What type of theoretical framework can be used to guide researchers and practitioners in this area?			different categories of stakeholders are identified with subsequent development of performance measurement and differentiated strategy to meet their needs and expectations. The use of a simplified performance prism was found to be effective in identifying and categorizing the stakeholders of the organisation. Staff at all levels had an understanding of the new system and perceived it as being beneficial. However, there were concerns that the approach was not continuously managed throughout the year and was in danger of becoming an annual event, rather than an on-going process. Furthermore, the change process seemed to have advanced without corresponding changes to appraisal and reward and recognition systems. Thus, the business objectives were not aligned with motivating factors within the organisation.
McAdam, O'Neill	2002		What is the benefit of clustered benchmarking in UK local government?		The clustering approach is more beneficial than comparing the service of a single unit with best practice (in the building control services).
Meier and O'Toole	2003	What difference does network-focused public management make for implementation? Do the ways that managers deal with complex surroundings make a difference in how programs work? How can managers make use of their			The core idea is that management matters in policy implementation, but its impact is often nonlinear. One way that public managers can make a difference is by leveraging resources and buffering constraints in the program context. This investigation finds empirical support for key elements of the network-management portion of the model. The results of this analysis suggest that network management matters even more in high-performing and low-performing cases.

		complicated and interdependent settings to enhance performance?			
Meier, O'Toole, Boyne, and Walker	2007		What is the influence of strategy content on organisational performance?		The results show that strategy can be separated out from other elements of management for a distinguishable assessment of its impact on organisational performance. The authors find that the defender strategy is the most effective for the primary mission of the organisation and that the prospector and reactor strategies work best in regard to the goals of the more politically powerful elements of the organisation's environment (related to four general types of strategic actors: prospectors, defenders, analysers, and reactors).
Melkers and Willoughby	2005			What are the effects of performance measurement information on budgetary decision making, communication, and other operations of U.S. local governments?	Research findings indicate the consistent, active integration of measures throughout the budget process is important in determining real budget and communication effects in local governments. The implementation of performance measurement supports improved communication within and across branches of government, advances learned discussion about the results of government activities and services, and adds value to budgeting decisions by providing relevant information about results, as well as costs and activities.
Micheli and Kennerley	2005		What are the requirements of a framework in order to allow public and non-profit organisations to better evaluate their own		The paper demonstrates that there are a number of key considerations when designing or choosing a framework for use in the public or non-profit sector: 1. It is necessary to understand the

			performance by means of collected data?		analogies and differences between public, non-profit and private sectors; 2. It will be necessary to identify all the stakeholders involved in public and non-profit organisations; 3. The main constituencies of the model and cause-and-effect relationships between them should be identified; 4. A framework should be sufficiently complex and comprehensive to include all the main features of the organisation, but at the same time comprehensible and flexible, enabling modification by those who will work with it.
Modell	2001			How do the properties of institutional processes associated with public sector reforms impinge on the extent of pro-active choice exercised by senior management in the development of PM?	The case findings yield several important insights pertaining to the research question and Oliver's (1991) hypotheses. The authors find evidence of both legitimacy-seeking and efficiency-enhancing rationales forming part of senior management's rhetoric.
Modell	2004			How do competing performance measurement myths impinge on organisational action?	The authors incline to the position that myths pivoting around the supremacy of goal-directed, multidimensional PM models, such as the Balanced Scorecard, may gradually replace the myth that public service provision may be improved by heavy reliance on financial control and come to affect operating-level action.
Modell	2009		What institutional research has been done so far on PMM in the public sector accounting literature?		Research is beginning to move beyond simplistic portrayals of PMM as decoupled or loosely coupled practices to pay greater attention to how it is implicated in the formative stages of

					institutional processes and the role of agency exercised at different levels of analysis. This signifies a shift in analytical foci from a one-sided emphasis on the institutional effects *on* PMM, treating institutional pressures as largely exogenous, to examine the more intricate roles of PMM as an outcome of as well as a medium for change.
Moriarty and Smallman	2009	What is the epistemology of benchmarking and which are the methodological elements of a theory of benchmarking?			Benchmarking remains theoretically underdetermined, with publications focusing on pragmatism and praxis rather than epistemology. Analysis of the literature leads to a new definition of benchmarking focusing around the teleological processes that lead to state-transformation of organisations. It is concluded that a theoretical foundation for benchmarking should be consistent with current organisational paradigms and the nature of what constitutes current and superior states of affairs.
Moynihan	2005	What are the insights of relevant organisational learning literature on results-based reforms?			From an organisational learning perspective, most results-based reforms target narrow process improvement (single-loop learning) rather than a broad understanding of policy choices and effectiveness (double-loop learning), even though the latter is more critical for long-term organisational success. Case evidence from state governments illustrates single- and double-loop learning and the importance of two frequently neglected aspects of organisational learning: learning forums—routines where performance information is deliberately examined - and the role of organisational culture in enabling or limiting learning.

Moynihan	2006			How did the implementation of public management reform in the United States take place?	It is argued that the managing for results doctrine has been only partially adopted. State governments selected some of the New Public Management ideas but largely ignored others. In short, state governments emphasised strategic planning and performance measurement but were less successful in implementing reforms that would enhance managerial authority, undermining the logic that promised high performance improvements.
Moynihan and Ingraham	2004			What is the role of integrative leadership in managing for results (MFR)?	This article argues that that one mode of effective leadership is active engagement in management systems, in this case, MFR. Leaders who become publicly involved in, and identified with, setting strategic goals and demanding performance information have a positive effect on encouraging employees to take MFR seriously as a process, and to incorporate performance information into decision making. Perhaps the most striking result of the analysis is how the leadership of different actor's matters in distinct ways for different types of decisions and the audiences associated with those decisions. The evidence suggests, therefore, that the influence of leadership will vary with the type of leader promoting the system, and the level of employee responding to leader initiatives. The finding suggests that if MFR systems turn into data-production compliance exercises this will actually discourage decision makers to use information. Size of government proves to be positively related to performance information use, and significant for senior executive branch decisions.
Mite	2000	Can performance management (PM) as a			The paper proposes that the adoption of the PM model is a universal remedy for improving service

		systems-based model cultivate the ``achievement culture'' in public sector organisations (PSOs) in developing countries?			quality in PSOs in developing countries. It is recommended that: 1. For the PM model to function effectively, performance of employees must be linked to corporate objectives, measured, and recognised. The planning, decision making and control processes must be based on performance information from the management accounting system; 2. There is a need to link PM systems design with issues of policy, strategy, operations, assessments and information systems; 3. Management accounting and other performance measurement practices need to be evaluated not just from an economic perspective, but also from a social, behavioural and managerial perspective, within an overall organisational context.
Newcomer	2007			What are the challenges and opportunities that performance measurement present for public managers (in the US government)?	A lack of clarity in expectations among public managers regarding how performance measures may be used presents a significant constraint inhibiting the effective use of such measures to inform managerial decision-making. Experience in a variety of government jurisdictions has shown that managers face complex communication, analytical, political, and measurement challenges in designing and implementing performance measurement systems. Yet despite these challenges, opportunities for using performance measurement to improve public management are endless, and the momentum carrying performance measurement forward seems quite strong.
Nicholson-	2006			What is the effect of a	The choices managers make about measurement

Crotty, Theobald and Nicholson-Crotty				manager's choice of performance measures on the assessment of organisational performance and decisions regarding solutions?	can have a significant impact on their evaluations, including assessments of (1) whether their organisation has a problem, (2) the environmental and organisational causes of the problem, and (3) whether their solutions to the problem are working.
Norman	2002		What are the lessons from experience with New Zealand's public sector performance management systems?		Research with a cross-section of users of this system, now nearly 15 years old, reveals a variety of responses. *True Believers* support a current focus on measurement and think that more effort should be put into creating clearer, more observable measures that emphasise outcomes. *Pragmatic Sceptics* see reported measures as part of a new game of public management and at best a starting point for asking about the substance behind the form. *Active Doubters* believe that too much emphasis on measurement gets in the way of the 'real work' of developing relationship-based work in a political environment. Issues of meaning are seen to be more important than measurement for the further development of the system.
Pandey and Garnett	2006	What is the effect of communication performance to agency effectiveness?			The findings of this research study have two key implications for public managers. First, the constraints of red tape on communication performance can be overcome if key performance-enhancing conditions - goal clarity without rigidity and a culture that supports communication - are in place. Second, external communication poses more challenges and may require additional effort.
Parker and	2000	Does organisational			Prescriptions for organisational practice derived

Bradley		culture in Queensland public sector organisations reflect the emphasis of the new public management on group, developmental and rational cultures?			from management and public choice theory are not presently being modelled in Queensland' public sector departments. Instead, culture in public sector organisations continues to reflect traditional approaches to public administration. The authors suggest that public sector organisations continue to emphasise the values of a bureaucratic or hierarchical organisational culture.
Pidd	2005			Why can performance measurement systems in public services lead to dysfunctional consequences even when people operate with the best of intentions?	Though many reasons are cited for public service performance measurement regimes, it is clear that control aspects dominate the others. This, when allied to an unthinking use of cybernetic metaphors, is what can lead to dis-functionality.
Poister and Streib	1999			To which extent has performance measurement been integrated into contemporary local government management in the United States?	Improvements were cited in a number of areas, but relatively few substantial effects were claimed. Big impacts are limited to behavioural change in terms of improved managerial accountability and increased employee focus on organisational goals, with much less frequent impact cited in terms of changes in program focus or priorities, budget allocations, cost savings or employee motivation.
Pollanen	2005			For what services, and to what extent, have efficiency and effectiveness measures been developed? For what purposes, and to what extent, are (should) efficiency and effectiveness measures	For every purpose, and for both efficiency and effectiveness measures, greater use was perceived desirable than actually occurred, and a significant increase in the use was expected in the near future for both types of measures, particularly for effectiveness measures. Performance measures were used most often for internal purposes, such as program management decisions, budgeting and resource allocation, comparing actual performance against targeted performance, and

				(be) used? What factors, and to what extent, impede the development, use, and reporting of efficiency and effectiveness measures?	reporting to elected officials. The following factors impede the development, use, and reporting of efficiency and effectiveness measures: • Difficulty in identifying appropriate measures; • Difficulty in meaningful use of measures; • Ambiguity of performance objectives
Pollitt	2005			How and to what extent do performance indicators influence the top management of the agencies concerned (in Finland, the Netherlands, Sweden and the United Kingdom) and the degree to which performance data were used by ministries as steering instruments?	In north western Europe, performance measurement has become almost universal. This is the level on which convergence is most evident, and it goes well beyond rhetoric and into practice. Performance management is growing steadily but varies in form and force among different countries and different tasks. Performance steering of agencies by ministries is still a rarity and may never become particularly common. Performance measurement and performance management remain activities conducted chiefly by and for managers. On the whole politicians do not take much interest, and neither do citizens—unless and until disasters, scandals, or breakdowns come along.
Pollitt	2006			What is already known of the use of performance information by ministers, parliamentarians and citizens (the end users)?	A literature review indicates that research into use by these groups has been very patchy, and that much of what we do know suggests that evaluations and performance reports and audits are seldom highly valued by politicians or citizens. Possible reasons and remedies for this apparent state of affairs are discussed on page 49 to 51.
Rantanen,	2007		What are the specific		The design and implementation processes in the

Kulmala, Lönnqvist and Kujansivu			problems faced by the Finnish public sector organisations in designing and implementing performance measurement systems (PMS)?		Finnish public sector organisations differ significantly from the way they are realised in industrial private sector companies. The four underlying reasons for problems in public sector organisations are the following: 1. There are many stakeholders with conflicting needs; 2. The end products and goals are undefined; 3. There is a lack of property ownership; and 4. Lacking management skills.
Rondeaux	2006	Is an identity evolution taking place following the implementation of NPM principles in the Belgian federal service?			The paper confirms the hypothesis that an identities evolution is taking place following the implementation of new public management (NPM) principles in the Belgian federal service. Using two identity logics “public service,” relating to the principles and values associated with traditional public administration, and “public managerialism”, relating to the principles, values and representations linked to NPM, it distinguishes six identity profiles. These are characterised by their positioning according to two principles fidelity and reality. The conclusion is that organisational identity is complex, hybrid and composite and in constant evolution according to perceptions of reality and context.
Rondo-Brovetto and Saliterer	2007		Can local government benchmarking help to enhance and consolidate regional and local economic performance and competitiveness?		Although the use of benchmarking has become a popular tool within public management reform, the authors conclude that the use of benchmarking as an instrument to improve local government performance only focuses on operational and organisational aspects. Current benchmarking strategies in the public sector are therefore unlikely to have a significant influence on improving services at the local level and

					therefore also make no contribution to the enhancement or sustainability of the competitiveness and performance of communities and regions.
Sanger	2008		To what degree are governments measuring and sharing the results citizens want, and how are they determining what those are? Further, to what degree do governments and agencies actually realise the potential of measurement for managing their operations, motivating and engaging their workforces, and allocating their resources?		Four principle lessons have emerged from this research: a. Performance measurement is growing in states and local governments, but more often without the engagement of citizens and with unrealised use for management; b. Some improvement is evident at all levels of government, but cities do better than states and performance management efforts are growing most successfully at the agency level; c. Jurisdictions and agencies with the best performance reporting and performance management efforts have strong mission driven leaders at the helm who communicate the mission, motivate employees, shape strategies, and provide support, rewards, and sanctions for achievement. d. The state of knowledge about what jurisdictions and agencies are doing, why, and with what success is growing but remains inadequate to inform intervention or policy.
Sharifuddin bin Syed-Ikhsan and Rowland	2004		What are the perceptions on the benefits, problems, responsibilities and technological aspects that are entailed in managing		a. Benefits: this study shows that knowledge management as a practice could be the most influential strategy in managing knowledge in public organisations in Malaysia in the near future b. Problems: the study revealed that the most

			knowledge in a public sector organisation in Malaysia?		difficult issue to manage was changing employees' behaviour c. Responsibilities: develop an organisational database of information and knowledge, develop effective and efficient methods of gathering information and knowledge, have systematic training for all employees and develop a culture that can promote knowledge sharing d. Technological aspects: e-mail, online information sources and the Internet are seen as very important
Streib and Poister	1999		What are the validity, legitimacy, and functionality of municipal performance measures?		*Validity:* There are many US municipalities struggling to develop meaningful PM systems. Many municipalities also face problems when compiling, analysing, and distributing PM data. *Legitimacy:* the authors learned that lower level employees are typically not involved in the development of performance measures. They also know that some municipalities experience resistance from both managers and lower level employees. Finally, they have learned that citizen involvement in the development of performance measures is a rare event. *Functionality:* The benefits of municipal performance measurement appear limited to manager accountability and employee focus on organisational goals. It does not appear that performance measurement regularly leads to changes in the focus or emphasis of programs, cost reductions, program priorities, budget allocations, or employee motivation.
Talbot	2008	What are the main groups			Evidence suggests that it is the totality of a

		of institutional actors who can attempt to shape or steer the performance of service delivery agencies?			performance regime which potentially shapes or steers performance for specific organisations rather than the narrow purchaser-provider or principal-agent assumptions often made about performance drivers.
Teelken	2008	What are the difficulties experienced with implementing performance measurement in the Dutch higher education and healthcare sector?			Institutional and organisational theory (combined into three dimensions: individual, group and organisational level) supports the argument that current features of performance measurement systems (PMS) in public organisations are generally unsuitable for the actual nature of these professional organisations. Despite external pressures, the implementation of such systems is slower than intended and seems to occur outside the primary process of the organisation. Institutional as well as professional theories supplement each other in a fruitful way in order to explain the difficulties with implementation of PMS. The 'human' side of PMS should be able to acknowledge and bridge the gap between the primary process and the organisation of performance measurement, which is obviously present in public, professional organisations. While institutional theory emphasises the initial resistance to change, professional theory helps to understand the pragmatic embrace by the individual.
Tillema	2007			What are the characteristics of Dutch Water Boards that may affect those organisations' use of benchmarking information for performance	The author concludes the following: 1. With a low exposure to economic markets, public sector organisations may ignore information (including benchmarking information) that indicates that their relative performance is poor;

				improvement? What are the mechanisms through which these characteristics function, and under which circumstances can these mechanisms be expected?	2. Public sector organisations may have difficulties imitating superior business processes. For this reason, public sector organisations might consider institutionalising the imitation process, which would imply that cooperation among the benchmarking partners is not only embedded in the performance measurement and analysis stage of a benchmarking project (as it was in the wastewater treatment benchmarking project), but also in the performance improvement stage; 3. A benchmarking culture and having powerful managers is of great importance.
Triantafillou	2007			What are the effects of benchmarking in the public sector?	The author adopted a conceptual framework that makes room for an analysis that suggests that the most important danger of benchmarking may be neither its 'perverse' effects nor its abuse by central authorities to retain control. The real danger of benchmarking may rest with the fact that even in the situation where the benchmarking analysis is initiated, designed and conducted by the benchmarked organisations themselves, they are caught in a game of incessant organisational change that some of the participants will always lose. As a device of power, benchmarking depends upon the production of normalizing knowledge and the freedom or self-governing capacities of those who are benchmarked.
Van Bockel and	2006	In what ways do			The authors argue the following:

Noordegraaf		instruments that introduce personal stakes in public domains "professionalise" collective action, aimed at improving the public good?			1. Performance-driven, NPM instruments are not merely about instrumental change, but about changes in identities that surround public organisations; 2. NPM not only affects instruments, but also affects the meaning of public organizing – and it does so in a biased way; 3. The NPM era is a next step in a long-term process of a professionalisation of public management, preceded by what the authors describe as "pre-Weberian" and "Weberian" eras; 4. "Post-Weberian" organisations individualise public organizing and institutionalise personal stakes, without strengthening a renewed sense of the *res publica*.
Van Dooren	2005	What makes organisations measure?			Six organisational factors have been studied (Belgium): 1. *Measurability* of the services of the organisations is a key factor for implementation. Organisations that have more routine-based services have a higher implementation and adoption of PM; 2. The degree of *political interest* for measurement does not explain the degree of adoption and implementation; 3. *Scale* is also relevant. Large organisations measure more. This invokes questions about the minimal capacity that organisations need to measure; 4. *Street-level discretion* – with the existence of time registration as an indicator – correlates positively with implementation but not with adoption; 5. The potential impediment of the lack of

					resources does not explain either implementation or adoption. The provision of sufficient resources is a critical factor for starting up as well as maintaining and extending performance measurement; 6. The *linkage between goals and indicators* seems to be of particular importance for the implementation of performance measurement. The decoupling does not seem to impede adoption. A PM policy is often the missing link in public sector reform. Insight in the organisational and contextual factors that facilitate or impede PM is crucial for developing a performance measurement policy. Governments may be tempted to design a one-size-fits-all policy, often based on the best practices in the public sector. However, differences between organisations may be considerable and should be taken into account.
Van Helden and Tillema	2005			What are the various response patterns of public sector organisations to benchmarking (based on insights from economic, neo-institutional and resource dependence theories)?	The authors argue the following: 1. The influence of various factors on the willingness of public sector organisations to engage in a benchmarking project is more univocal than their influence on the willingness to consider performance improvement actions as a result of benchmarking information; 2. Some factors that influence the willingness of organisations to take part in a benchmarking project and the resulting performance improvement actions may conflict with others; 3. It is questionable whether uncertainty stimulates organisations to use

					benchmarking information for performance improvement.
Van Helden, Johnsen and Vakkuri	2008		What is the nature of public sector performance measurement research?		The authors show that accounting researchers from Europe investigate reasons for limited PM use and factors explaining a rational or symbolic PM use, inspired by organisation theory and institutional theory and conducting case/field studies. Public administration researchers from Europe and the USA prefer to study PM design and PM impact respectively, mainly using surveys in combination with various theories, like political theory. Public administration research from the USA examines the types of performance indicators in PM systems and contingent factors for PM design. Public administration research from Europe shows an interest in evaluating public sector reforms like Best Value and explaining learning processes for improvement. The authors argue that PSPM research could benefit from interdisciplinary efforts and intensified mutual communication between public administration and accounting.
Van Thiel and Leeuw	2002			How to deal with the performance paradox in the public sector?	The performance paradox refers to a weak correlation between performance indicators and performance itself. The increase in performance assessment in the public sector following the administrative reforms of the 1980s and 1990s has had several unintended consequences, threatening insight into performance and performance itself. To counteract these consequences, performance assessment systems should take the special characteristics of the public sector into account. The contested nature of performance indicators requires the use of

					multiple indicators, referring to different aspects of policy implementation (tangible and non-tangible) and reflecting the interests of all stakeholders (politicians, managers, funders, providers, purchasers, and consumers). Moreover, a balance has to be found between too much and not enough measure pressure.
Vigoda-Gadot and Yuval	2003		What is the relationship between managerial quality, administrative performance and citizen's trust in government and in public administration systems in Israel?		The study shows that managerial quality leads to administrative performance and ultimately to trust in governance. The findings support previous assumptions that administrative performance may be treated as a precondition to trust in governance rather than trust serving as the precondition to performance.
Willcocks	2002		What is meant with public sector managerial effectiveness?		Managerial effectiveness is essentially about understanding, reinterpreting and making sense of different role expectations, for which the author offers a framework. The conceptual framework proposed focuses upon different levels of public sector effectiveness – in particular, individual, managerial, organisational, and inter-organisational levels of effectiveness. Given the complexity, ambiguity, and subjective nature of the concept of effectiveness in the public sector, it has attempted to argue the case for multi-theoretical and a multi-level, conceptual framework.
Williams	2004		What has been the development of performance measurement in the critical period from its origins through 1930?		PM originated at the early Bureau of Municipal Research (after 1906). Over the next quarter century, it became more sophisticated through increased quantification and reliance on experts. However, its focus narrowed from government to government service. This narrowing is linked to

					reduced social activism among those who used these methods. The entire period saw combined interest in accomplishing results and containing costs. This history also shows us that *performance measurement* does not refer to a particular empirical technique. Instead, it refers to the application of relevant techniques to the problem of observing government at work; that meant the delivery of government services.
Wisniewski and Stewart	2004			What are the performance information needs of diverse local authorities in Scottish local authorities? To what purpose will performance information be put by those using it?	Performance information needs are service specific, and therefore there is unlikely to be a set of common performance measures across different services. Few of the pilot sites of this study had given much thought as to the most appropriate way of reporting performance and had done little to assess stakeholder satisfaction with either the performance information provided or the way it had been provided.
Wynn-Williams	2005			How can benchmarking help to provide meaningful and relevant information to funders, service providers, service recipients and other interested parties?	The study proposes that a combination of internal benchmarking, process benchmarking and increased public documentation will enhance reporting systems in any public sector organisation.
Yang and Hsieh	2007	How do the unique characteristics of the public sector affect the dynamics of performance measurement in Taipei (Taiwan)?			The results indicate that the implementation of performance measurement is inseparable from the evolution of politics and democratic governance, suggesting that integrating political science constructs (e.g., political support) and organisational theory constructs (e.g., organisational support) can better explain the public management phenomenon. Future studies

					should continue in this direction, integrating performance measurement with other political science and organisational theory constructs. The practical lesson from this study is that in order to institutionalise performance measurement and make it work, public managers must ensure top management commitment, middle manager support, stakeholder involvement, continuous training, and external political support.
Yasin	2002	What are the gaps with theoretical and practical implications?			The literature related to benchmarking practices and theory was reviewed from 1986 to 2000. The earlier stages of benchmarking developments stressed a process and/or activity orientation. Recently, however, the scope of benchmarking appears to have expanded to include strategies and systems. Despite recent advancements, the field of benchmarking still suffers from the lack of theoretical developments which are badly needed to guide its multi-faceted applications.

Appendix 3 Classification of the studies: focus and theoretical orientation

Author	Date	Focus 1 Measurement	Focus 2 Incorporation	Focus 3 Use	Orientation	Method I	Method II	Approach
Adcroft and Willis	2005		incorporation	use	positivist	obtrusive small N		single case study
Alstete	2008		incorporation	use	interpretative	obtrusive small N		online discussion board and e-mail correspondence
Amaral and Sousa	2009		incorporation		positivist	obtrusive small N		single case study
Ammons, Coe and Lombardo	2001		incorporation		positivist	obtrusive large N		survey
Ammons and Rivenbark	2008			use	positivist	obtrusive small N		survey and face-to-face interview
Anand and Kodali	2008		incorporation	use	positivist			literature study
Andersen, Henriksen, and Spjelkavik	2008		incorporation	use	positivist	obtrusive small N		action research and multiple case study
Andrews, Boyne and Walker	2006		incorporation		positivist	obtrusive large N		survey
Askim, Johnsen and	2007			use	positivist	obtrusive large N		survey

Cristopherson								
Ball, Bowerman, Hawksworth	2000			use	positivist	obtrusive small N		face-to-face interview
Behn	2003	measurement			positivist			literature study
Berman and Wang	2000			use	positivist	obtrusive large N		survey
Bogt, ter	2004			use	positivist	obtrusive large N		survey
Bogt, ter	2008 (February)		incorporation	use	positivist	obtrusive small N		interviews
Bogt, ter	2008 (August)		incorporation	use	positivist	obtrusive large N		semi-structured interviews, document analysis
Boland and Fowler	2000		incorporation		positivist	obtrusive small N		multiple case study
Bowerman, Francis, Ball and Fry	2002			use	positivist	obtrusive small N	obtrusive large N	face-to-face interview and survey
Braadbaart	2007			use	positivist	obtrusive small N		quasi-experimental face-to-face interview and survey
Braadbaart and	2008	measurement	incorporation	use	positivist			literature study

Yusnandarshah								
Brignall and Ballantine	2004			use	positivist			literature study
Brignall and Modell	2000		incorporation	use	positivist			literature study
Broadbent and Laughlin	2009		incorporation	use	positivist			literature study
Brown, Waterhouse and Flynn	2003		incorporation		positivist	obtrusive small N		single case study
Cavalluzzo and Ittner	2004		incorporation	use	positivist	obtrusive large N		survey
Coplin, Merget and Bourdeaux	2002			use	positivist			literature study
Dattakumar and Jagadeesh	2003	measurement	incorporation	use	positivist			literature study
Dawes, Cresswell and Pardo	2009		incorporation		positivist	obtrusive small N		action research
De Bruijn	2002		incorporation	use	positivist			literature study
De Bruijn and Van Helden	2006		incorporation	use	positivist	obtrusive small N		multiple case study
Fernandez and Rainey	2006		incorporation	use	positivist			literature study
Ferreira and	2009		incorporation	use	positivist	obtrusive small		multiple case

Otley						N		study
Folz	2004	measurement			positivist	obtrusive large N		survey
Foster, Gallup	2002	measurement			interpretative	obtrusive small N		telephone survey
Fryer, Antony and Ogden	2009		incorporation	use	positivist			literature study
Garnett, Marlowe and Pandey	2008		incorporation		positivist	obtrusive large N		survey
Goncharuk and Monat	2009	measurement	incorporation		positivist	obtrusive small N		longitudinal field study
Greiling	2005			use	positivist			literature study
Grubnic and Woods	2009	measurement			positivist			literature study
Hinton, Francis, Holloway	2000		incorporation		positivist	obtrusive large N		mailed survey
Holloway, Francis and Hinton	1999			use	positivist	obtrusive large N	obtrusive small N	survey and single case study
Holzer and Kloby	2005		incorporation	use	positivist			literature study
Holzer and Yang	2004		incorporation	use	positivist	obtrusive small N		multiple case study
Horton	2006		incorporation	use	positivist			literature study

Jansen	2008			use	positivist	obtrusive small N		multiple case study
Johnsen	2005		incorporation	use	positivist			literature study
Johnsen and Vakkuri	2006		incorporation	use	positivist			literature study
Jones	1999		incorporation		positivist	obtrusive small N		single case
Julnes and Holzer	2001			use	positivist	obtrusive large N		survey
Kouzmin, Löffler, Klages and Korac-Kakabadse	1999		incorporation	use	positivist	unobtrusive large N		document analysis
Kyrö	2003	measurement			interpretative			literature study
Kyrö	2004	measurement			interpretative			literature study
Laise	2004	measurement						literature study
Longbottom	2000		incorporation		positivist	obtrusive large N	obtrusive small N	mailed survey and face-to-face interview
Magd and Curry	2003			use	positivist	obtrusive small N		multiple case study
McAdam, Hazlett and Casey	2005	measurement			interpretative	obtrusive small N		single case study, face-to-face interviews and focus groups

McAdam, O'Neill	2002		incorporation		positivist	obtrusive small N		focus group, mailed survey and face-to-face interview
Meier and O'Toole	2003		incorporation	use	positivist			literature study
Meier, O'Toole, Boyne, and Walker	2007			use	positivist	obtrusive large N		survey
Melkers and Willoughby	2005		incorporation	use	positivist	obtrusive large N		survey
Micheli and Kennerley	2005	measurement			positivist			literature study
Modell	2001			use	positivist	obtrusive small N		case study
Modell	2004		incorporation	use	positivist			literature study
Modell	2009	measurement	incorporation	use	positivist			literature study
Moriarty and Smallman	2009	measurement	incorporation	use	positivist	obtrusive large N		survey
Moynihan	2005		incorporation	use	positivist	obtrusive small N		multiple case study
Moynihan	2006		incorporation		positivist	obtrusive large N		survey
Moynihan and Ingraham	2004			use	positivist	obtrusive large N		survey
Mite	2000			use	positivist			literature study

Newcomer	2007		incorporation	use	positivist			literature study
Nicholson-Crotty, Theobald and Nicholson-Crotty	2006	measurement			positivist	obtrusive large N		survey
Norman	2002		incorporation	use	positivist	obtrusive large N		interviews and survey
Pandey and Garnett	2006		incorporation	use	positivist			literature study
Parker and Bradley	2000		incorporation		interpretative	obtrusive large N		survey
Pidd	2005			use	positivist			literature study
Poister and Streib	1999		incorporation		positivist	obtrusive large N		survey
Pollanen	2005			use	positivist	obtrusive large N		survey
Pollitt	2005			use	positivist	obtrusive small N		document analysis and semi-structured interviews
Pollitt	2006			use	positivist			literature review
Rantanen, Kulmala, Lönnqvist and Kujansivu	2007	measurement			positivist	obtrusive small N		multiple case study
Rondeaux	2006		incorporation		interpretative	unobtrusive	obtrusive small	document analysis (content

						large N	N	and speech analysis), face-to-face interviews (contextual and identity),
Rondo-Brovetto and Saliterer	2007			use	positivist			literature study
Sanger	2008			use	positivist	obtrusive small N	unobtrusive large N	face-to-face interview and document analysis
Sharifuddin bin Syed-Ikhsan and Rowland	2004		incorporation		positivist	obtrusive small N	obtrusive large N	single case study and questionnaires
Streib and Poister	1999			use	positivist	obtrusive large N		survey
Talbot	2008		incorporation	use	positivist			literature study
Teelken	2008		incorporation	use	positivist	obtrusive small N		multiple case study
Tillema	2007			use	positivist	obtrusive small N		multiple case study
Triantafillou	2007			use	positivist	obtrusive small N		multiple case study
Van Bockel and Noordegraaf	2006		incorporation		positivist			literature study and historical analysis

Van Dooren	2005		incorporation	use	positivist	obtrusive small N		survey, face-to-face interviews
Van Helden and Tillema	2005		incorporation	use		obtrusive large N	obtrusive small N	survey and multiple case study
Van Helden, Johnsen and Vakkuri	2008	measurement	incorporation	use	positivist			literature study
Van Thiel and Leeuw	2002			use	positivist			literature study
Vigoda-Gadot and Yuval	2003		incorporation		positivist	obtrusive large N		survey
Willcocks	2002		incorporation		positivist			literature study
Williams	2004	measurement			positivist			literature study
Wisniewski and Stewart	2004	measurement			interpretative	obtrusive small N		multiple case-studies
Wynn-Williams	2005			use	positivist	unobtrusive large N		document analysis and literature study
Yang and Hsieh	2007		Incorporation	use	positivist	obtrusive small N		single case study and survey
Yasin	2002	measurement			positivist			literature study

Appendix 4 List of interviewees

Unstructured interviews (12) conducted between December 2008 and September 2009

Organisation	Date interview	Function
VNG (these persons were all transferred to KING in 2009)	05-12-08	Three project managers
VNG	14-01-09	Director
Zenc	10-02-09	Partner Zenc responsible for the implementation of the Building and Housing Supervision benchmark
SenterNovem	11-02-09	Two advisors
SGBO	05-03-09	Advisor
VNG (these persons were all transferred to KING in 2009)	05-03-09	Three project managers
KING	25-03-09	Interim Director KING
Zenc	15-04-09	Partner Zenc responsible for the implementation of the Building and Housing Supervision benchmark
Association BWT	27-05-09	Chairperson Association BWT
Breda	31-08-09	Deputy Head Public Affairs Division
Groningen	25-09-09	Head Building and Housing Department
Emmen	28-09-09	Team Leader Supervision

Semi-structured interviews (24) conducted between November 2009 and September 2010

Municipality	Date interview	Function	Level
Breda	23-11-09	Head Public Affairs Division Deputy Head Public Affairs Division	Strategic
	18-01-10	Section Head (Centre)	Managerial
	18-01-10	Coordinator Supervision and Enforcement	Operational
	18-01-10	Coordinator Inspectors	Operational
	19-01-10	Section Head (West)	Managerial
	19-01-10	Section Head (East)	Managerial
	19-01-10	Coordinator Construction and Supervision	Operational
	24-02-10	Head Public Affairs Division Deputy Head Public Affairs Division	Strategic
	24-02-10	Alderman Building and Housing	Strategic
Eindhoven	22-06-10	Head Permits	Managerial
	30-09-10	Executive officer permits, supervision and enforcement	Operational
	30-09-10	Senior Policy Officer	Managerial
	30-09-10	Team Leader Review and Enforcement	Managerial
Emmen	08-12-09	Head Permits	Managerial
	08-12-09	Section Head Building	Managerial
	15-12-09	Alderman Building and Housing	Strategic
	15-12-09	Team Leader Supervision	Managerial
	15-12-09	Head Supervision	Managerial
	15-12-09	IT officer (supply of data to be used for the benchmark questionnaire)	Operational

Groningen	06-01-10	Team Leader Building and Housing	Managerial
	06-01-10	Team Leader Policy and Support	Managerial
	06-01-10	Alderman Building and Building	Strategic
	29-04-10	Quality Officer	Operational
	29-04-10	Head Building and Housing Department	Managerial

Appendix 5 Results Benchmark Building and Housing Supervision 2009

Perspective	Issues covered	Description	Breda	Eindhoven	Emmen	Groningen
Client perspective	Fees	Percentage of the total revenues relative to the total rated construction volume. The lower the fees are, the better it is for the applicant.	1,59%	1,31%	0,88%	2,69%
	Turnaround	The indicator is the duration of a license request with a deadline of 12 weeks. The shorter the duration, the better.	5,4 weeks	5 weeks	12 weeks	3,3 weeks
	Draft design	The number of preliminary designs as a percentage of total applications. Here, the higher the better.	27%	22%	15%	27%
	Inadmissible	The percentage of applications that is ultimately declared inadmissible. The lower the rate, the fewer applicants need to be disappointed, the better it is.	8,7%	10,4%	5,5%	9,9%
	Direct admissible	The percentage of requests that can be treated without the applicant having to provide additional data. The higher this percentage, the better the information and / or intake worked for both the municipality and applicant.	56%	45,3%	59,1%	56,6%
	Opening hours	Number of front office opening hours per week. The more hours the municipality is open, the better it is for the applicants.	36 hours	35 hours	42,5 hours	40 hours
	Front office capacity	Percentage of total fee required to cover formation that is used for front office tasks and information.	7,2%	9,3%	3,6%	7,8%
Management perspective	Accuracy	The percentage of appeals and objections relative to the total number of decisions taken. The lower the better.	1.0%	6,5%	2,4%	0,2%
	Quantity of policies	Number of aspects related to building and housing on which the municipality has adopted concrete testable policy. It is assumed that the more policy, the better the municipality has its act together.	16	24	22	17
	Statutory deadlines	The percentage of requests that lead by law to authorisation relative to the total number of applications. The lower the percentage the better the municipality controls its processes.	1,7%	2,8%	1,0%	1,3

	Coverage	The extent to which the fee income covers the costs to be covered by the fees. The higher the cost coverage, the better.	136,1%	82,2%	99,5%	138,9%
	Formation of enforcement	Amount of capacity that the department has available for carrying out enforcement (in particular illegal construction) as a percentage of the total section formation.	17,3%	7,0%	6,2%	13,1%
	Special procedures	Percentage of all building permits applications that need to follow a special procedure. The lower the rate, the better the municipality's seems to have her zoning plan in order.	36,1%	25,5%	21,2%	28,8%
	Quality level of the review	The score of the index test intensity is plotted against the minimum acceptable standard as agreed jointly by the municipalities in the LTP norm. The higher the score above the LTP-norm, the more intensive review is done than regarded as minimum acceptable.	103	70	89	84
Organisation perspective	Absenteeism	Percentage of total staff that could not work productive in 2008 due to illness. The lower the rate, the better it is.	6,3%	5,4%	2,9%	4,4%
	Education level	Percentage of employees related to total formation with a degree in higher education. It is assumed that the higher the percentage, the better the quality.	58%	84%	8%	84%
	Loss control	Total percentage of requests that ultimately do not result in a permit. The lower this percentage, the more efficient people work.	20%	35%	11%	29%
	Average cost per building volume	The height of the costs is shown as the cost per € 10,000 assessed building volume. The lower the cost per unit rated construction volume, the better it is.	€ 117	€ 159	€ 89	€ 194
	Average cost per FTE	Average costs per FTE to cover from construction fees (including costs third parties, overhead and direct costs).	€ 130,694	€ 158,921	€ 140,031	€ 165,870
	Average cost per application	Prefix indicating the average cost for the organisation to treat an application. The lower the cost per application, the better.	€ 3,228	€ 4,533	€ 2,666	€ 4,159
	Reviewed construction volume	Productivity measure wherein the total amount reviewed building volume is divided by the formation that may be covered from the construction fees. The more building volume is assessed per FTE, the higher the productivity, the better it is.	€ 11,181,087	€ 10,002,193	€ 13,052,153	€ 9,850,722

	Productivity (no. of applications per FTE)	Prefix that covers the number of handled requests per FTE to be covered from the construction fees. The more applications a FTE treats on average, the better it is.	40,2	36,7	43,5	44,9

Source: Perspectives and issues covered in the questionnaire of the Building and Housing Supervision benchmark (source: Zenc (October 2008), Proposal Benchmark Building and Housing Supervision Round 2009, The Hague), Zenc (October 2009), *Benchmark Building and Housing Supervision Round 2009 (about the performance from the year 2008): graphs of all municipalities*, The Hague.