



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Aki Jääskeläinen

**Productivity Measurement and Management in Large
Public Service Organizations**



Julkaisu 927 • Publication 927

Tampere 2010

Aki Jääskeläinen

Productivity Measurement and Management in Large Public Service Organizations

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Konetalo Building, Auditorium K1702, at Tampere University of Technology, on the 26th of November 2010, at 12 noon.

ISBN 978-952-15-2459-2 (printed)
ISBN 978-952-15-2469-1 (PDF)
ISSN 1459-2045

ABSTRACT

Jääskeläinen, Aki. 2010. *“Productivity Measurement and Management in Large Public Service Organizations”*. Department of Industrial Management. Tampere University of Technology, Tampere, Finland.

Keywords: measure, measurement, performance, performance management, productivity, public sector, public service

Productivity in public services is always a topical theme. There is a constant need to pay attention to maintaining and improving productivity. These pressures have increased due to many current challenges in the Finnish public sector. There are various ways to improve productivity at many levels of examination. Much of the existing research on the topic has been carried out at the macro level. This study examines productivity from the perspective of management in individual public departments and units. Measurement is seen as an essential tool supporting productivity improvement. The main aim of this research is to investigate the development of measurement systems supporting the productivity management of large public organizations. The research is loosely based on the commonly used phasing of developing performance measurement systems. It has two main research questions: 1) how can public service productivity be described as a phenomenon? 2) how can productivity management be supported by means of measurement? There are many research themes underlying this study, the most essential of which relates to public sector performance measurement and management.

This dissertation, as a whole, can be characterized as a qualitative multiple case study. It is based on six research articles applying various research methods. Interviews, action research and a statistical study were the main methods. The empirical context is the City of Helsinki, Finland, which is a large municipal organization providing and arranging a wide range of public services. The research questions are examined organizationally at the top (e.g. municipal department) and bottom (e.g. unit providing a specific service) levels. However, the main emphasis is in the operative level examination, which is carried out in the context of social services.

The main contribution of this research relates to the description of a bottom-up measurement approach supporting productivity management in large public organizations. The main idea is first to develop appropriate component measures for the operative level, after which the results from each measure can be aggregated for purposes of top-level decision-making. The choice of this approach is supported by the existing literature and prior experiences from practice. The initial experiences of the approach are positive in the organization studied. No prior studies describing empirical application of similar approach were found in the literature. Other contributions of this research relate to two issues. First, new knowledge is provided on the productivity phenomenon in the context of public services. A new model for the comprehensive examination of factors affecting productivity is presented. Second, this study contributes to the existing research by describing the process of developing a measurement system in an interesting specific setting. Two key factors affecting the success of the development process are identified: the commitment of the operative level and the suitability of the measurement system for the requirements of the organization.

TIIVISTELMÄ

Julkisten palveluiden tuottavuus on aina ajankohtainen teema. Tuottavuuden ylläpitoon ja parantamiseen on jatkuva tarve kiinnittää huomiota. Suomen julkisella sektorilla näitä paineita lisää tällä hetkellä monet haasteet. Voidaan tunnistaa monia tapoja parantaa tuottavuutta riippuen siitä, mitä tarkastelutasoa käytetään. Suuri osa olemassa olevasta aiheeseen liittyvästä tutkimuksesta on tehty makrotasolla. Tämä tutkimus tarkastelee asiaa julkisten virastojen ja julkisia palveluita tuottavien yksiköiden johtamisen näkökulmasta. Mittaaminen nähdään eräänä keskeisenä tuottavuuden parantamisen apuvälineenä. Tutkimuksen pää tavoite on tarkastella tuottavuuden johtamista tukevien mittausjärjestelmien kehittämistä suurissa julkisissa organisaatioissa. Tutkimuksen rakenne liittyy usein käytettyyn mittausjärjestelmien kehittämisprosessin vaiheistukseen. Tutkimus muodostuu kahdesta pää tutkimuskysymyksestä: 1) miten julkisten palveluiden tuottavuutta voidaan kuvata ilmiönä? 2) miten tuottavuuden johtamista voidaan tukea mittauksen keinoin? Tutkimuksen taustalla on monia tutkimusteemoja joista tärkein liittyy julkisen sektorin tuloksellisuuden mittaamiseen ja johtamiseen.

Kokonaisuudessaan tätä väitöstutkimusta voidaan luonnehtia laadulliseksi, useista tarkastelukohteista muodostuvaksi tapaustutkimukseksi. Tutkimus muodostuu kuudesta artikkelista, joissa on hyödynnetty erilaisia tutkimusmenetelmiä. Päämenetelminä on käytetty haastatteluja, toimintatutkimuksia ja tilastollista tutkimusta. Empiirisenä tarkastelukohteena on Helsingin kaupunki, suuri kunnallinen organisaatio, joka tuottaa ja tilaa suuren joukon erilaisia julkisia palveluita. Tutkimuskysymyksiä tarkastellaan organisatorisesti sekä ylä- (esim. virasto) että alatasoilla (esim. yksittäinen palveluita tuottava yksikkö). Suurin huomio on kuitenkin operatiivisen tason tarkastelussa, joka toteutetaan sosiaalipalveluiden kontekstissa.

Tutkimuksen pääkontribuutio liittyy suurten julkisten organisaatioiden tuottavuuden johtamista tukevan ”bottom-up” mittauslähestymistavan kuvaukseen. Siinä kehitetään soveltuvia komponenttimittareita aluksi operatiiviselle tasolle, jonka jälkeen kunkin mittarin tulokset voidaan yhdistää ylätason päätöksenteon tarpeisiin. Tätä lähestymistapaa tukee aikaisempi kirjallisuus ja käytännön kokemukset. Ensimmäiset kokemukset lähestymistavasta ovat positiivisia tämän tutkimuksen kohdeorganisaatiossa. Tiedossa ei ole aikaisempia tutkimuksia, jotka olisivat empiirisesti soveltaneet vastaavaa lähestymistapaa. Tutkimuksen muu kontribuutio voidaan tiivistää kahteen teemaan. Ensinnäkin, tutkimus tuottaa uutta tietoa tuottavuusilmiöstä julkisissa palveluissa. Se esittää uuden mallin, jota voidaan hyödyntää tuottavuuteen liittyvien tekijöiden kattavassa tarkastelussa. Toiseksi, tutkimus kuvaa mittausjärjestelmien kehittämisprosessia kiinnostavassa erityistapauksessa. Tutkimuksessa tunnistettiin kaksi keskeistä tekijää, jotka edesauttavat mittausjärjestelmien kehittämisessä onnistumista: operatiivisen tason edustajien sitouttaminen ja järjestelmän soveltuvuus organisaation tarpeisiin.

ACKNOWLEDGEMENTS

Writing a doctoral dissertation is no pushover, nor should it be. This research work is a result of many phases, feelings of happiness and frustration. In fact, it was only my third topic that eventually led to a concrete end result. This was fundamentally facilitated by a suitable research project during which it was possible to study one topic in-depth. Throughout my life, I have tried to successfully finish every major effort I start. That is why I am very happy to do so with this personal research work. In retrospect, these four years with the topic of this research actually feel surprisingly systematic and straightforward. This would not have been the case without many people whom I now wish to thank.

First, I wish to thank the supervisors of this research. The guidance and support from Professor Petri Suomala were important during the last and important phases of my dissertation work. Help was always forthcoming when needed, but I also had freedom to implement my own decisions. Professor Erkki Uusi-Rauva was my supervisor during the first three years of this research. Erkki was also the supervisor of my master's thesis. I greatly appreciate his experienced guidance and feedback, which has been an important cornerstone for this research. I am proud to continue the research stream that Erkki started many years ago.

I am very grateful to Professor Antti Lönnqvist, who has had a major role in all the phases of my journey as a researcher. Antti taught me a lot about the work of the researcher. I appreciate his valuable feedback, support and co-operation throughout my research work. My sincere gratitude goes also to the pre-examiners of my dissertation, Professor Hannu Rantanen (Lappeenranta University of Technology) and Adjunct Professor Juha Kostiainen (Sitra) for their professional comments and remarks.

During all my time as a researcher since March 2005, I have been a part of the Performance Management Team at Tampere University of Technology. I want to thank all the past and present members of the group for providing support and a superb working environment. I appreciate the valuable contribution by Dr. Paula Kujansivu and Virpi Sillanpää M.Sc. in the research project and articles related to this thesis. Special thanks go also to Dr. Harri Laihonen, Jonna Käpylä M.Sc. and Maiju Vuolle M.Sc. The support of the Department of Industrial Management has been important for me. Special thanks go to Jukka Annala and Sirpa Järvenpää for all the help in daily work. I am also very thankful to Virginia Mattila for the English language proofreading of this work even within my tightest schedules.

This research was carried out in close contact with the City of Helsinki. I am grateful for the opportunity to carry out research in such a prominent organization. The support from various personnel as well as the opportunity to gather rich research data were very important for me. Especially I want to thank chief planning officer Jukka Öberg and financial manager Leena Karhu-Westman for comments and other support during the various phases of research and development work.

It is impossible to carry out research without sufficient financial support. My dissertation work was financed by the Finnish Work Environment Fund, the City of Helsinki, the Finnish Doctoral Program in Industrial Engineering and Management and Tampereen Liikesivistyssäätiö. I am grateful for the much needed support.

I have always wanted to separate work and leisure. I greatly appreciate the support from my family as well as all the friends. During my leisure time, it has always been easy to forget all the worries regarding my research work. Finally, I want to thank Piia for her patience and support even during the most frustrating periods in the journey towards this achievement.

Tampere, October 6th 2010

Aki Jääskeläinen

TABLE OF CONTENTS

ABSTRACT	i
TIIVISTELMÄ	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	v

PART 1: INTRODUCTORY ESSAY1

1. INTRODUCTION2

1.1 Background	2
1.2 Key concepts	3
1.2.1 Productivity and related concepts	3
1.2.2 Concepts related to measurement in organizations	9
1.2.3 Performance and productivity management	11
1.2.4 Public service	12
1.3 The productivity phenomenon in the public sector	14
1.4 Productivity measurement and management of public services	19
1.4.1 Phases and challenges in measurement	19
1.4.2 Managerial and other uses of measurement	23
1.4.3 Measurement approaches and methods	25

2 RESEARCH DESIGN31

2.1 Research gap, objective and questions	31
2.2 Scope of the research	34
2.3 Description of the empirical context	35
2.4 Research methodology	41
2.4.1 Research approach	41
2.4.2 Data and methods	43
2.5 Research structure	49

3 RESULTS54

3.1 How can public service productivity be described as a phenomenon?	54
3.1.1 What affects productivity at different organizational levels?	54
3.1.2 What is the impact of different productivity factors at the operative level?	60
3.2 How can productivity management be supported by means of measurement?	61
3.2.1 What kinds of needs for measurement can be identified at different organizational levels?	61
3.2.2 How can the needs be satisfied with measures?	64

3.2.3 What affects success in the process of developing measurement systems?	71
3.3 Summing up the results.....	73
4 CONCLUSIONS	76
4.1 Contribution to the existing research	76
4.2 Contribution to practice	78
4.3 Assessment of the research	79
4.3.1 Relevance	79
4.3.2 Validity.....	79
4.3.3 Reliability.....	82
4.3.4 Generalization	84
4.4 Suggestions for further research	85
REFERENCES	88
APPENDICES	103
 PART 2: PUBLICATIONS	 105
I Jääskeläinen, A. (2010), Identifying factors affecting public service productivity, <i>International Journal of Services Technology and Management</i> , Vol. 14, No. 4, pp. 360–375.	
II Jääskeläinen, A., Lönnqvist, A., Public Service Productivity: How to Capture Outputs? <i>International Journal of Public Sector Management</i> (accepted for publication).	
III Jääskeläinen, A., Kujansivu, P., Lönnqvist, A. (2009), Learning from Existing Performance Information, <i>Proceedings of 5th Conference on Performance Measurement and Management Control, September 23-25, 2009, Nice, France</i> .	
IV Jääskeläinen, A., Uusi-Rauva, E., Bottom-up Approach for Productivity Measurement in Large Public Organizations, <i>International Journal of Productivity and Performance Management</i> (accepted for publication).	
V Jääskeläinen, A. (2009), Identifying a Suitable Approach for Measuring and Managing Public Service Productivity, <i>Electronic Journal of Knowledge Management</i> , Vol. 7, No. 4, pp. 447–458.	
VI Jääskeläinen, A., Sillanpää, V. (2010), Evaluating the factors of success in the performance measurement system development of two public services, <i>Proceedings of 10th EURAM Conference, May 19-22, 2010, Rome, Italy</i> .	

PART 1: INTRODUCTORY ESSAY

1. INTRODUCTION

1.1 Background

In general, it is easy to find motivation for a study investigating productivity. Productivity improvement has often been linked to the wealth and competitiveness of nations (Craig and Harris, 1973; Sink, 1983). It also seems to have a positive effect on the wages of the labor force, employment and living standards (Mammone, 1980). Overall wage increases should be linked to productivity improvement in order to sustain the balance of the national economy (Hjerppe and Kangasharju, 2003, p. 11). Productivity is important whatever the industry or sector. The specific interest of this study is in the Finnish public sector. Productivity is a topical theme in the public sectors of many countries. Not surprisingly, this is also the case in Finland where a substantial number of public services is provided compared to many other countries (Hautakangas et al., 2007). The public sector accounts for around a fifth of the Finnish economy, meaning that it cannot be ignored when discussing productivity (Hautakangas and Heikkinen, 2008).

According to the calculations of Statistics Finland (2008), the productivity of Finnish local administration has decreased by around 1-3 percent almost every year in the 21st century. Many challenges cause pressures to improve productivity in the Finnish public sector. A key challenge relates to the aging of the population, which is also an issue in many other western countries. It has been estimated that the retirement of employees has the most substantial impacts in the public sector (Hovila and Okkonen, 2006). In Finland, around half of the employees in health and social services will retire by 2015. In addition, there will be more customers for those services. Individual customers may also become increasingly demanding to serve. (Halinen and Korhonen, 2008, p. 22) It has been estimated that there will be a need for an increase of 200,000 employees in public welfare services by 2040 (in comparison to year 2007) if the service providing structures remain unchanged (Parkkinen, 2007). If the level of taxes is not substantially raised, the productivity of public services must be improved (Halinen and Korhonen, 2008, p. 21). Another key challenge relates to the ongoing economic recession, meaning that there are limited financial resources available for public service provision. If productivity is not at good level, tax-money is not optimally used to satisfy the needs of the public (cf. Rosen, 1993, p. 2). Productivity can therefore also be linked to sustaining desired outcomes in the public sector (such as public health) and even the very existence of many public services.

In many countries, criticism of the inefficiency and ineffectiveness of the public sector led to a transformation already in the 1980s and 1990s. Many changes, such as decentralization and privatization, have been made as a part of this transformation. (Van Helden, 2005) Such dramatic changes are not the only way to improve productivity. At the level of organizations, productivity is often regarded as an essential component of organizational performance (Tangen, 2005) indicating the need for attention in daily management. Along with the many changes in the public sector, new management techniques from the private sector have been applied as a part of New Public Management (NPM) (Hood, 1995; Pollitt and Summa, 1997). Public sector performance measurement has been an important issue for decades, but its importance has been increased by the New Public Management movement (Greiling, 2005; Hood, 1995; Johnsen, 2005; Kloot and

Martin, 2000; Sanderson, 2001; Van Helden et al., 2008). Performance measurement has been mandated in the public sectors of several countries (Johnsen, 2005). Public organizations are required to demonstrate that there have been improvements in performance and that goals and objectives are being achieved (Wisniewski & Stewart, 2004). In Finland, the Ministry of Finance requires the state administration to implement performance measures in order to set targets and manage performance (Salminen and Viitala, 2006). Functional measurement can provide many benefits in organizations, such as improved decision relevance, identification of problems and successes, increased accountability and transparency (Johnsen, 2005). In practice, however, many challenges and problems have been identified in measures used by public organizations. Models and frameworks from the private sector are possibly not applicable as such in the public sector (Pidd, 2005; Radnor and McGuire, 2004; Wisniewski and Stewart, 2004).

In the context of many public services, evaluation and measurement have focused on performance aspects other than productivity (Laine, 2005). However, public organizations also need measurement information related to the productivity of their service production. This information is useful in demonstrating productivity to society and tax-payers. Productivity measures can also be utilized in improving productivity by identifying concrete targets for development. In general, productivity measurement is an old topic, especially at the level of industries and nations, and in the context of the manufacturing industry (Singh et al., 2000). Productivity measurement in the public sector has been criticized for not capturing the unique characteristics of services. Some general challenges have been related to inaccurate data and poor output measures (e.g. Kangasharju, 2008, p. 212; Ministry of Finance, 2007). If measures are poor, there is a risk that means to improve productivity will be inappropriate resulting in possibly negative consequences for the productivity of services. One possible reason for deficient measures is that complexity of productivity phenomenon in the context of public services is not well understood. Another problem may be that productivity measures are not integrated into the operative management of public organizations indicating poor linkage to the general knowledge on performance measurement and management.

This dissertation aims to investigate the development of measurement systems supporting the productivity management of large public organizations. The empirical examination, carried out as a qualitative multiple-case study, was conducted in the context of a large Finnish municipal organization, the City of Helsinki. The thesis comprises an introductory essay and six scientific articles. The essay consists of four chapters: 1. Introduction, presenting the key concepts and theoretical background, 2. Research design, illustrating the research gap and questions, likewise the methodology used, 3. Results, describing key findings in relation to the research questions posed, 4. Conclusions, presenting the contribution and an evaluation of the research.

1.2 Key concepts

1.2.1 Productivity and related concepts

There are several concepts the meanings of which are similar to productivity. There also seems to be various and even conflicting views related to the definition of the concepts (Rautiainen, 2004). This is an inevitable cause of confusion when discussing the topic and also of challenges in

measuring productivity. In this section, the concepts productivity, quality, profitability, efficiency, effectiveness and performance are discussed in detail. It is also presented how these concepts are understood in this study.

Productivity

The origin of productivity examination is related to the context of industrial manufacturing and agriculture (Grönroos and Ojasalo, 2004, Uusi-Rauva, 1997, p. 16). Productivity is usually defined as the ratio between output (e.g. quantity of products and services produced) and input (e.g. labor, material, capital) (Chew, 1988; Craig and Harris, 1973; Sink, 1983). Most productivity models and definitions examine the efficiency of a production process either directly or indirectly (Hannula, 1999, p. 2). According to Tangen (2005), productivity is related to the use of resources and productivity decreases if resources are utilized inefficiently. Since productivity is also related to outputs, it includes the element of value creation. On the other hand, waste may be considered to be the opposite to what productivity symbolizes. In the context of industrial production, productivity is often related to physical phenomena, meaning that outputs and inputs represent physical units (Banker et al., 1989; Hannula, 1999). Consequently, productivity is not increased merely by applying higher prices to products nor does it decrease due to higher costs caused by inflation. This is related to the difference between profitability and productivity, which is examined in more detail later in this section.

There are various opinions on quality in relation to the productivity concept. Grönroos and Ojasalo (2004) claim that there has traditionally been an assumption of consistent quality in the productivity examination. According to Tangen (2005), improvements in quality should not be included in the concept of productivity. The only exception is that defective products should not be considered as outputs. According to Hannula (1999, p. 31), the relationship between quality and productivity depends on the definition of both concepts. The total quality of an organization or a process is supportive of productivity. He also states that product quality should be taken into account in productivity measurement. Quantifying quality changes is therefore a measurement problem not a conceptual problem. Quality examination has been emphasized in studies on service productivity (Parasuraman, 2002; Rosen, 1993). One reason may be the fact that in many services all the outputs (also 'defective') are received by customers. Aspects related to service productivity are discussed later in more detail.

Quality

The concept of quality is often used at a general level without paying too much attention to defining it. It is a challenging concept which is difficult to define (Folz, 2004; Gummesson, 2000, p. 157). It has even been argued that there can be no single and all-embracing definition for quality since quality-related characteristics vary depending on the object of examination (Gaster, 1996). Quality is often related to satisfying customer needs. According to Gummesson (1998), quality may be defined as doing things right from the beginning and doing those things that customers need and want. In this study, the specific interest is in examination of services. Rosen (1993, pp. 56 - 57) contends that service quality can be determined by its usefulness to the client, and continues that a public service has better quality if it is more accurate, prompt, durable, reliable, convenient,

accessible and courteous. In practice, service quality can be improved, e.g. by paying attention to the control of the service providing capacity in order to avoid queuing of customers (Sherwood, 1994). In addition, factors such opening hours and geographical location can improve the availability of services and at the same time the quality perceived by customers. Some factors related to service quality, such as reliability, responsiveness and empathy are clearly intangible in nature and therefore difficult to control (Wakefield and Blodgett, 1999).

Hence, several aspects of service quality can be distinguished, which makes it difficult to operationalize the concept. Grönroos (2001) identifies three general dimensions in the quality of services:

1. Technical quality refers to the output or outcome of a service which can often be measured in a fairly objective manner
2. Functional quality is related to the actual process of service production (e.g. fluency of service provision)
3. Perceived service quality is a function of perceived and expected quality. This refers to the fact that the customer often evaluates service quality subjectively in relation to pre-expectations.

According to Gummesson (1998), the productivity discussion has traditionally paid attention to the technical quality aspects related to the design of products and production. However, increasing attention is paid to quality perceived by customers which may be measured, e.g. by customer satisfaction surveys.

Profitability

Even though the main interest of this study is in the public sector, in which profitability is often not a meaningful concept due to lack of information related to output prices, it is examined briefly here in order to achieve a more comprehensive picture of productivity-related concepts. Profitability can be defined as the ratio between revenue and costs (Tangen, 2005). Profitability is the key driver of success in every private industry and takes special account of the needs of shareholders. However, there is increasing criticism of the short-term and retrospective perspective of profitability examination (Gummesson, 1998). It has long been known that productivity appears to be a better measure for examining the long-term excellence of production (Miller, 1984). In spoken language productivity and profitability are sometimes used synonymously. This may be one reason why companies forget the significance of productivity. (Tangen, 2005)

Since productivity and profitability are closely related concepts, it is important to make a distinction between them. Productivity is an essential factor affecting the profitability and competitiveness of organizations (Hannula, 1999; Rantanen, 1995). However, the connection between productivity and profitability is not always unambiguous (see Figure 1). Profitability is affected by changes in productivity and price recovery. Changes in revenue are therefore the result of both changes in product quantity and price. Similarly, costs are affected by changes in resource quantity and price. Hence, there is not always a connection between profitability and productivity. Inflation and other

external market-based factors may affect profitability even if there are no changes in productivity (Stainer, 1997).

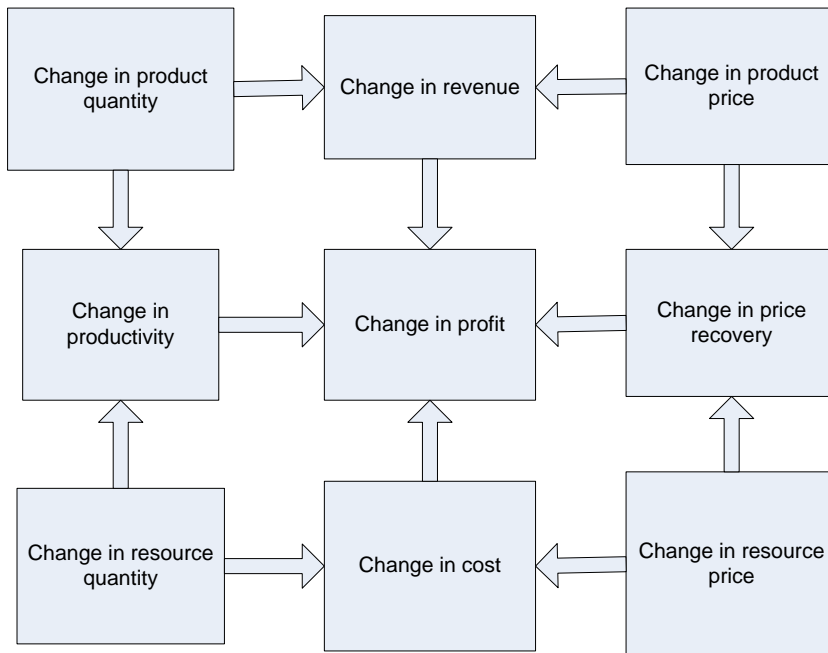


Figure 1 Factors affecting profitability (adapted from Loggerenberg ja Cucchiaro, 1982).

Efficiency

Efficiency is a concept closely related to productivity and has been defined in various ways. It may be defined as the ratio between realized and standard or expected production (Florentin et al., 1991, p. 132; Hannula, 1999, p. 29). It can also be defined simply as doing things right (Drucker, 1963). Rosen (1993, p. 93) defines efficiency in the following way: efficiency is related to output/input examination when output quality is ignored. Rautiainen (2004, p. 34), on the other hand, sees efficiency as the relationship between provided benefits and resources. Despite the differences in definitions it seems that most researchers relate efficiency to the use of resources and therefore to the nominator of the output/input ratio (Tangen, 2005). Hence, efficiency is defined as the minimum resource level theoretically required to maintain the desired operations compared to the actual level of resource consumption. In this conception efficiency is closely related to utilization rate.

Effectiveness

Effectiveness is a concept that has sometimes been confused with efficiency (Hannula, 1999, p. 29). Effectiveness may be defined simply as doing the right things (Drucker, 1963). According to Hannula (1999, p. 29), effectiveness is related to the external performance of a process whereas efficiency to the internal performance. Effectiveness may be defined as the ability to reach a desired objective or the degree to which the desired results are achieved. Hence, effectiveness seems to be related to the ability to produce high product quality. According to Neely et al. (1995), effectiveness is related especially to satisfying customer needs. Similarly, Tangen (2005) relates effectiveness to

value creation from the customer's perspective. Effectiveness is therefore more closely related to the numerator of the output/input ratio.

The concept of effectiveness has been emphasized in the public sector, where organizations' key objective is related to the welfare and health of the general public instead of profits. Rosen (1993, p. 51) distinguishes public program effectiveness and the effectiveness of the implementation process. Program effectiveness is related to the achievement of desired outcomes by having a positive impact on the situation or problem at which it was targeted. It is related to both choice and implementation of means. Implementation effectiveness is related to the actualization of a public program, the extent to which the mandated services have been produced and delivered. According to Rosen, implementation effectiveness, which may be measured by multiplying output quantity and output quality, is most relevant when discussing productivity in the public sector.

Performance

According to Thomas and Baron (1994), there is a tendency to extend the productivity discussion in a way that the term performance would be more appropriate. Kaydos (1999) regards operational performance as the efficiency and effectiveness of business processes. Performance may also be defined as the ability of an organization to achieve defined objectives (Institute of Industrial Engineers, 1990, pp. 11–14). According to Tangen (2005), performance is related to almost any objective of competition and manufacturing excellence. It can be seen as an umbrella term for all the concepts that examine the success of an organization and its activities. The perspectives of performance may be very case-specific.

Organizational performance is usually related to multidimensional examination taking into account all the key stakeholders of an organization. For example, the much used Balanced Scorecard for performance measurement includes perspectives of finance, internal processes, customers as well as learning and growth (Kaplan and Norton, 1992). Slack et al. (2001) examines operational performance from the perspectives of low costs, flexibility, speed, reliability and quality. According to Sink (1983), the overall performance of a company includes at least seven criteria:

- efficiency
- effectiveness
- quality
- productivity
- quality of work life
- innovations
- profitability.

Finally, performance may be examined from three different aspects (Lönqvist, 2004, p. 28). First, performance is related to the results or outputs of certain activities. Second, performance is used to refer to doing an activity. Third, performance is linked to the ability to achieve results.

Linkages between the concepts

Figure 2 summarizes the conceptual examination of this section. At the same time it illustrates how the concepts are understood in this research. Efficiency is related to the utilization of inputs and doing things right. Productivity, in turn, examines the output of a production process including quantity and quality of products and services. Quality is related specifically to the examination of outputs. Quality is regarded as a part of service output and also as an essential link between the outputs and outcomes of public services. Effectiveness may be connected to the outcomes and benefits which are examined in relation to the organization's objectives and customer needs. Outcomes are partly the results of outputs but may be also affected by other factors (e.g. a customer). Effectiveness should be high when an organization is pursuing the correct objectives in an efficient manner. Costs are affected by the quantity and prices of inputs used. Revenue is the result of both the quantity and prices of products and services sold. Revenue is dependent on both outputs and outcomes from operations. Profitability is related to the relationship between revenue and costs. Finally, performance is a broad concept which includes all the sub-concepts of the figure.

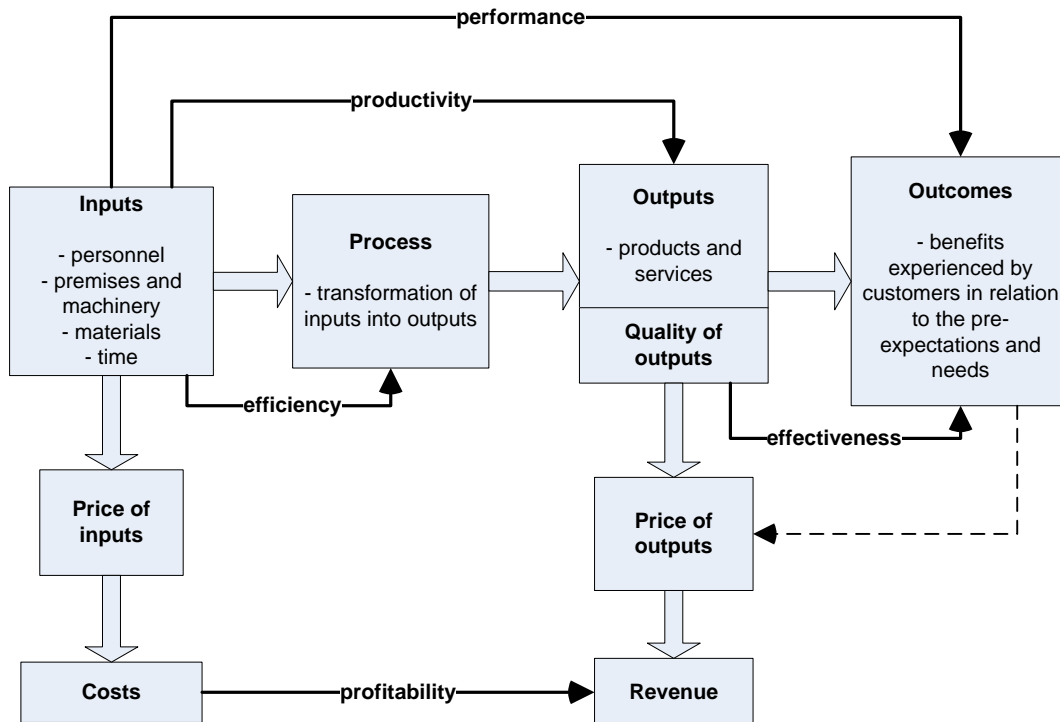


Figure 2 Linkages between productivity and related concepts.

Despite the seemingly clear definitions of the concepts, their operationalization is not always straightforward. For example, productivity measurement in a certain setting requires context-specific evaluation.

1.2.2 Concepts related to measurement in organizations

Measurement

Bunge (1973, p. 120) defines measurement as an effective determination of the value of a magnitude or quantity with the help of empirical operations (Vehmanen, 2008). According to Kaydos (1999, p. 15), measurement consists of assigning a numerical scale to the size, value or other characteristic of a tangible or intangible object. A key aspect in measurement is the measurement scale used. There are four main types of scales in variables (Bryman and Bell, 2007, pp. 355 - 356):

- Interval/ratio variables have identical distances between categories across the range of categories. Celsius temperature is an example of a measurement scale providing such variables. Ratio variables are similar to interval variables but they have a fixed zero point. Ratio scales are used, for example, in measuring the costs of operations.
- Ordinal variables have an order of categories but the distances between the categories are not the same across the scale. Scales (e.g. totally agree, agree, disagree, totally disagree) used in questionnaires often provide ordinal variables.
- Nominal variables, also known as categorical variables, include categories that cannot be rank ordered. Examples of such categories could be different nationalities of people.
- Dichotomous variables (e.g. gender) contain only two categories and they have therefore only one interval. They may have attributes of the other three types of variables.

The term measurement can be used for different purposes. The more detailed meaning of measurement is therefore related to the measurement object and purpose of measurement. In the discussion of public sector productivity, Rosen (1993, p. 2) regards measurement as a way to provide solid information that public officials require when making rational and defensible decisions about the allocation of resources. Kaydos (1999, p. 19) regards measurement as a way of providing reliable and meaningful information for managers. Similarly in this research, measurement, in general, is seen as a tool for providing information on the object considered managerially relevant. Measurement may be used as a part of various managerial activities, not only in a retrospective manner but also in planning and forecasting. The managerial use of measurement is discussed more closely in Section 1.4.2. Since the research on performance measurement links measurement and management in a practical way, this study also relies on the definition of performance measurement. Neely et al. (1996, p. 11) define performance measurement as the process of quantifying the efficiency and effectiveness of purposeful action. Lönnqvist (2004, p. 31) has presented a definition of performance measurement which is suitable for different managerial purposes: “performance measurement is a process used to determine the status of an attribute or attributes of the measurement object”. This definition of measurement is also used in this study.

Measure and measurement object

Measurement object is a factor or phenomenon that is considered relevant enough to be measured. In the field of performance measurement, the measurement object is usually called a success factor (Lönnqvist, 2004, p. 32). This means that measurement objects are factors that drive the success of organizations, which is generally assured by somehow linking them with the organization’s mission and strategic objectives. Measurement objects may be related e.g. to the resources, activities,

outputs and outcomes of an organization. The difference between a measurement object and a measure should be acknowledged since they are sometimes confused. There may be various measures for a certain measurement object. This is very likely with complex measurement objects (such as the competence of employees) that are difficult to define.

Lönnqvist (2004, p. 33) defines a *measure* as the means for determining the status of an attribute or attributes of a measurement object. This general definition is also appropriate for the purposes of this study. Measures may be sorted in many ways (see e.g. Kaydos, 1999, p. 16; Simons, 2000, pp. 234–235) such as direct and indirect, financial and nonfinancial, subjective and objective measures. In practice there may be various tools (e.g. questionnaires) and formulas providing quantitative information on the aspects of a measurement object.

In this study, the specific interest is in productivity as a measurement object. Traditionally the measurement of productivity has been somehow related to the examination of the ratio between output and input. Such measures can be called direct measures of productivity. Indirect productivity measures, on the other hand, are related to factors (e.g. capacity utilization) considered to have an effect on productivity (see e.g. Uusi-Rauva, 1997, p. 67). Both direct and indirect measures should be considered in discussing productivity measurement with linkage to management. Productivity measures are commonly related to the group of nonfinancial measures (Hannula, 1999, p. 2) which is also the case in this study. Traditional output/input measures are clearly objective. In obtaining comprehensive information of complex and often intangible measurement objects, subjective measures are used (Lönnqvist, 2004). Subjective measures have been used for measuring knowledge work productivity (Antikainen, 2006). In this study, both objective and subjective measures related to productivity are examined.

Measurement approach, method and system

Some confusion may occur in the discussion on measurement techniques, tools and structures in organizations. In this study, the term *measurement approach* is used to refer to a general ‘philosophy’ of measurement which may be implemented using many different methods. *Measurement method* (e.g. formula) is defined as a way to provide information that can be used in productivity measures. (cf. Hannula, 1999, p. 50 - 51) Measures may be used as such e.g. as a part of statistics and reporting. However, it is common to use many measures that form a *measurement system*. Lönnqvist (2004, p. 33) defines a performance measurement system as a set of measures that determine the status of the attributes of the measurement objects. There are various frameworks, such as the Balanced Scorecard by Kaplan and Norton (1992), which may be used in developing performance measurement systems (frameworks related to public sector are discussed in Section 1.3). The use of such frameworks may at best help to avoid imbalance and separation of measures. In practice there may be several measurement systems for different organizational levels, different departments and organizational units.

Measures related to productivity may be used as a part of performance measurement systems. However, separate measurement systems focusing specifically on productivity can also be designed. In this study, the term productivity measurement system is used to describe a system consisting of several productivity measures used to compose a productivity index of productivity level or change.

Individual measures may be partial and indirect productivity measures or measures representing the productivity of separate organizational entities (productivity measurement is discussed in more detail in Section 1.4.3). The advantage of a separate system for productivity measurement can be that productivity gets specific attention and more effort is invested in gathering proper information on such an important measurement object (cf. Hodginson, 1999).

1.2.3 Performance and productivity management

Performance management is a term quite commonly used and may also be linked to productivity examination. Since performance management has many different applications depending on the purpose and the organizational level, there is, however, no established definition for the concept (Ukko, 2009, p. 5). For example, Williams (1991, p. 23) defines performance management as a process in which objectives are defined and monitored with measures in order to implement organizational strategy. Although this definition may be valid in the upper levels of an organization, it is not flexible when discussing different uses of measurement information. Hannula and Lönnqvist (2002) have presented a practical definition of performance management: it is management based on the information provided by performance measurement. In their view, performance management involves the systematic use of measurement for managing and developing the performance of different organizational activities.

The purpose of measurement is crucial in the more detailed meaning of the term performance management. The managerial use of measurement also differs at the operative and strategic levels. At the operative level, measures are used to provide detailed information supporting daily managerial activities. Objectives are set and progress is monitored by using measurement information and finally corrective actions are taken when necessary (Kaydos, 1999, p. 142; Kaplan and Norton, 1996, p. 251). This may be called single-loop learning (Argyris and Schön, 1978). At the strategic level, measurement is more focused on the key issues from the perspective of an organization's mission and strategic objectives. In this application, the strategy itself may be evaluated, e.g. by using measurement information for challenging predominant assumptions on cause-effect relationships (Kaplan and Norton, 1996, p. 267). This may be called double-loop learning (Argyris and Schön, 1978).

In the discussion on productivity, the concept of productivity management may be used. However, it is even less established in comparison to that of performance management. Sink (1985, p. 23) has stated that productivity management includes 1) measuring and evaluating productivity, 2) planning for productivity improvement and control, 3) making control and improvement interventions and 4) measuring and evaluating the impact of the interventions. Measurement has an essential role in this definition. Planning of productivity improvement and control should be done on the basis of productivity measurement. In addition, interventions related to productivity improvement should also be demonstrated with measurement information. Rosen (1993, pp. 240 - 241) has defined productivity management more broadly in the context of the public sector. However, the role of measurement is also emphasized in this definition. According to Rosen, managing for productivity consists of:

- Productivity measurement: the starting point for everything is measurement and even simple measures are better than impressions

- Managing the work: identifying sources of inefficiency in operations, modifying organizational structures, job design, work flow or technology (“engineering approach”)
- Managing the worker: ascertaining employee morale, considering training, reducing absenteeism and employee turnover etc. (“human relations approach”)
- Managing the management: productivity management involving appropriate use of all resources and managing for productivity meaning openness to the environment (e.g. client needs, alternative arrangements) and a constant search for better efficiency and better service quality.

The activities listed by Rosen are all relevant but there may also be many other more detailed tasks. One might ask: what is productivity management not? A broad definition of productivity can be related to everything that makes an organization function better (cf. Pritchard, 1995, p. 2). In addition, a lot of research has been carried out in the field of productivity without any intentional link to the productivity phenomenon (Käpylä et al., 2010). In this study, productivity management is taken to include all the managerial activities, with the specific and intentional aim of improving organizational productivity. With productivity as an ultimate objective, productivity management represents only a part of broader performance management. However, as discussed later in this study, the differentiation between the management of productivity and performance may be difficult and even unnecessary in practice.

1.2.4 Public service

There is no comprehensive and generally accepted definition for the concept of service. According to Hill (1977), a service may be defined as a change in the conditions of a customer or in goods belonging to the customer. Johnston and Clark (2008) regard service as the combination of outcomes (e.g. benefits and emotions) and experiences delivered to and received by a customer. When discussing service productivity, examination of service processes is important. According to Sampson and Froehle (2006) the presence of customer inputs is a necessary and sufficient condition to define a production process as a service process. Customer inputs may be related to physical presence, participation and the mind of a customer, material as well as information. In all these definitions, the role of customers is highlighted both from the point of view of service outputs/outcomes and inputs.

A traditional way of describing services has been the comparison between services and products. Four characteristics (“IHIP”) have been identified: intangibility, heterogeneity, inseparability and perishability (Grieves and Mathews 1997; Regan, 1963). Intangibility is related to the abstractness and non-physical nature of services. Heterogeneity means that it is often difficult to standardize services. Inseparability refers to the difficulty of separating service production from consumption. Finally, perishability is a characteristic which is related to the difficulty of storing services.

These characteristics have been criticized in the literature since they are not necessarily valid in all services (e.g. there are also standard services). In addition, some of the features, such as intangibility, may also be related to products. One reason for the difficulties in establishing an all encompassing definition and characterization of services is that the scope of services has been widened a lot in recent decades due to various innovations in digital business solutions (Viitamo,

2009, p. 10). For example, there are services in which consumption and production are separated (e.g. Internet banking). These novel commodities could also be regarded as intangible goods or hybrids (ibid.).

Services may be classified in many ways (Elinkeinoelämän keskusliitto, 2005; Lovelock, 1983; Silvestro et al., 1992). From the perspective of productivity, there are two classifications which are interesting specifically due to their impact on service provision. First, there are basically three possibilities in the role of customer in service provision: the service provider producing the service in isolation from the customer (back office), the service provider and the customer producing the service in interaction (service encounter) and the customer producing the service in isolation from the service provider (Grönroos and Ojasalo, 2004). Second, the service may be expert service with high customization or mass-produced service with a standardized content (Silvestro et al., 1992). The traditional characterizations of services are most valid with classic (or pure) services (cf. Gupta, 1995) in which there is close interaction between service provider and customer. They may also be described as high-touch services (Grönroos, 2000, p. 49) which refers to the employee intensity in service provision.

The public sector offers many critical services for the general public related to education, health, social welfare, security, water supply, transportation etc. Public services are provided or arranged by municipalities or the state. They are often provided in sectors in which there are no or few private companies operating, even though privatization has recently been an increasing trend. Some services, such as welfare services, are provided by the public sector since political control and decision-making are deemed important (Kangasharju, 2008, p. 195). Examples of such controlled questions are: who is in the greatest need of health care or what should be the content of certain services. Depending on the specific public service, there are many social objectives such as public welfare and health that are pursued.

According to Johnston and Clark (2008), public services have many interest groups and customers. It is not always clear who is the customer (client, relative, user, taxpayer etc.). The end user of public services is often someone other than the payer and therefore end users have only few options in choosing different services and service providers. Services are financed mainly from tax revenue and political decisions affect the allocation of resources between different services. Rosen (1993, p. 4) describes the operating environment of public organizations in comparison to private sector as follows: public organizations operate in more tightly constrained conditions since missions are fixed by law, operations are open to public evaluation and since rules and instructions define the options available of operating. In addition, civil service and budget systems limit freedom to redeploy labor and monetary resources. In discussing productivity, one key feature of public services is the lack of proper markets. There is often no price information on services and therefore sale values remain unknown (Hodgkinson, 1999, Kangasharju, 2008, p. 180).

In this study, the broad concept of public service is used to describe the empirical context and object of productivity measurement and management. There is a wide range of services in the public sector. Public organizations are often very heterogeneous enterprises (Edwards and Thomas, 2005). There are both mass (e.g. child day care) and expert services (e.g. surgical services). There is also

service production with almost no customer contact (e.g. urban planning), close interaction between service provider and customer (e.g. health care) and prominent customer role (infrastructure, energy and water supply etc.). Although it is acknowledged that there are various services provided by the public sector, this study aims to purposely use a more general approach in order to ensure the generalization of the results. However, the most detailed examination of this study is related to welfare services, which comprise educational, social and health care services (OECD definition, Elinkeinoelämän keskusliitto, 2005). These services form a more uniform group which could be described as classic services in which many of the traditional assumptions of the service literature apply fairly well.

1.3 The productivity phenomenon in the public sector

The purpose of this section is to present an overall picture of the perspectives related to the productivity phenomenon in the public sector. From the traditional productivity formula it follows that there are in principle five different circumstances in which productivity may be improved (Misterek et al., 1992):

- output increases faster than input
- more output from the same input
- more output with fewer inputs
- same output with fewer inputs
- output decreases while input decreases more

These circumstances can be achieved in various ways by managerial means. The actual transformation of inputs into outputs is often complex and may be affected by many factors. There are only few general models of public service productivity (e.g. Stainer and Stainer, 1998) while similar issues have been discussed more under the broader topic of public sector performance (e.g. Boyne, 2002; Talbot, 1999). On the other hand, productivity in services in general has been an issue of interest in several recent publications (Brax, 2007; Gummesson, 1998; Grönroos and Ojasalo, 2004; Johnston and Jones, 2004; Ojasalo, 2003; Parasuraman, 2002) in which a wider perspective on productivity has been applied. As an example, Vuorinen et al. (1998) define service productivity as an organization's ability to use its inputs in order to provide services satisfying the needs of customers. The contents of some key service productivity models are summarized in Table 1.

Table 1 Examples of aspects in service productivity models

Effect of service provider	Effect of customer	Author(s)
Company's perspective <ul style="list-style-type: none"> ▪ Output (profits and market share) ▪ Input (labor, equipment, technology) 	Customer's perspective <ul style="list-style-type: none"> ▪ Output (customer satisfaction) ▪ Input (time and effort needed) 	Parasuraman, 2002
Operational productivity <ul style="list-style-type: none"> ▪ Observable and quantifiable outputs such as number of customers and revenue ▪ Inputs such as materials, staff and costs 	Customer productivity <ul style="list-style-type: none"> ▪ Abstract outputs such as outcomes and values perceived by customer ▪ Inputs such as time and effort needed 	Johnston and Jones, 2004
	Service outputs <ul style="list-style-type: none"> ▪ Quality of customer resources ▪ Diversity of customer demands ▪ Ability to control who utilizes the service ▪ Customer competencies and willingness to co-produce 	Ojasalo, 2003
	Service provider's inputs <ul style="list-style-type: none"> ▪ Customer as a resource and co-producer ▪ Ambiguous role of customer, e.g. what kind of resources are offered and how? 	
<ul style="list-style-type: none"> ▪ Output quantity ▪ Service provider's inputs (personnel, technology, systems, time) 	<ul style="list-style-type: none"> ▪ Output quality ▪ Demand (factor affecting the utilization of service providing capacity) ▪ Customer inputs (own participation, participation of fellow customers) 	Grönroos and Ojasalo, 2004
<ul style="list-style-type: none"> ▪ Output quantity (volume) ▪ Input quantity (personnel, material, capital) ▪ Input quality (e.g. employee competence) 	<ul style="list-style-type: none"> ▪ Output quality perceived by a customer 	Vuorinen et al., 1998

Productivity has traditionally been related solely to the perspective of the provider of a product or service (Gummesson, 1998). This traditional conception of productivity has been criticized in the context of services (Gummesson, 1998; Grönroos and Ojasalo, 2004), where the perspective of customer has been emphasized (Ojasalo, 2003; Parasuraman, 2002). In classic services, a customer participates in the service provision and is therefore claimed to have a role in improving or impairing of the quality and productivity of services (Gummesson, 1998). Customers may affect both quantity and quality of outputs but also inputs (Ojasalo, 2003). In many services there are no outputs without customers since services cannot be stored (Gupta, 1995). Anticipating customer demand and the efficient use of service providing capacity is therefore essential in improving productivity (Grönroos and Ojasalo, 2004). Demand may also be affected by advising customers on the availability of service capacity (e.g. hours when there are fewer customers) (Rosen, 1993, p. 119).

Quality of services is another issue which has been stressed in connection with service productivity (e.g. Hodginson, 1999; Sahay, 2005). Parasuraman (2002) states that service quality influences outputs both for the company and the customer. According to Gummesson (2000, p. 158), there are challenges in controlling productivity of services, which is not at such a level as in manufacturing. In manufacturing, deficiencies may and should be identified before products are handed to the customers. This may not be so straightforward in service production. Grönroos and Ojasalo (2004) argue that since services are immaterial in nature, their observation and evaluation is mainly subjective. Therefore, productivity and perceived quality can even be seen as inseparable phenomena.

According to Johnston and Jones (2004), high productivity is generally good for both the provider and the customers in manufacturing since it means lower costs and prices. In service production, on the other hand, higher productivity by the provider may lead to lower productivity for the customer. High productivity of the service provider may have a negative impact on the quality perceived by customers and eventually also on profitability (Grönroos and Ojasalo, 2004).

The discussion on service productivity seems to relate especially to the classic services in which there is close interaction between customer and service provider. The background of the authors appears often to be in services marketing and the perspective of the customers is much emphasized. According to their conception, services not satisfying the needs of customers and not generating ideal outcomes have a negative impact on productivity. In the manufacturing context, productivity has been related only to the perspective of the producer, not the customer. However, it seems that the sharp contrast between manufacturing and services may not be meaningful since satisfying customer needs is equally important in manufacturing. Despite the wider perspective of service productivity models in relation to the productivity definition used in this study, they have many interesting aspects (related e.g. to quality of services) which are clearly important from the point of view of improving the productivity of public services.

It is difficult to find managerial models of public service productivity. Productivity in the public sector has often been related to cost-efficiency and quality of services (Faucett and Kleiner, 1994; Hodginson, 1999). Stainer and Stainer (1998) have presented a public service productivity model including inputs, processes, outputs and outcomes. They state that inputs are usually measured by costs, preferably in real terms. Outputs (physical terms) are the immediate result of the productive processes that ultimately affect the quality of life. The outcomes are the goals in social terms (e.g. better educated and healthier population). According to Rosen (1993, pp. 86 - 93), inputs of public services are labor, equipment, supplies, property and utilities. Labor is the major input in the public sector and represents most of the costs. Outputs, on the other hand, are related to the quantity and quality of services provided. Rosen specifies that quality of outputs means that difficulty in providing various services should be somehow taken into account.

Much of the literature related to managing productivity of public services has been published in the field of public sector performance. It has been noted in studies on public performance management that the frameworks from the private sector (such as the Balanced Scorecard) may not cover all the elements of organizational performance that are important to public organizations (Talbot, 1999).

Many different frameworks and classifications related to public service performance can be found in the literature, which are summarized in Table 2.

Table 2 Public sector performance models

Performance aspects	Model/Author
<ul style="list-style-type: none"> Effectiveness (cost-effectiveness and profitability) Customer Processes Performance ability of personnel 	Balanced Scorecard (adjusted to public organizations)/ Lumijärvi, 1999
Derived from mission instead of strategy	Balanced Scorecard (adjusted to non-profit organizations)/ Kaplan, 2001
<ul style="list-style-type: none"> Financial Customer Internal (processes) Learning and Growth 	
<ul style="list-style-type: none"> Effectiveness Efficiency (e.g. economy, productivity, profitability) Outputs and quality control (output volume, service ability and quality) Human resources development 	Tulosprisma/ Ministry of Finance, 2007
<ul style="list-style-type: none"> Outputs (quantity and quality) Efficiency (costs per unit of output) Outcomes (formal effectiveness and impact) Democratic outcomes (probity and participation) Responsiveness (satisfaction of staff and customers) 	Boyne, 2002
Enablers	Public Service Excellence model /
<ul style="list-style-type: none"> Strategic (policy and governance, strategy, leadership) Operational (resources, processes, people) 	Talbot, 1999
Organizational Results	
<ul style="list-style-type: none"> Internal results (resources, efficiency, people) External results (reporting, outputs, satisfaction) 	
Programme Results	
<ul style="list-style-type: none"> Outcomes Satisfaction 	
<ul style="list-style-type: none"> Economy (costs related to resources) Efficiency (outputs related to inputs) Effectiveness (achievement of objectives) 	3Es/ Midwinter, 1994 (originally in Accounts Commission, 1988)

Many of the classifications and models of public service performance have similar perspectives such as efficiency and effectiveness, even though they may emphasize certain specific aspects such as service quality or human resources. An important insight is that productivity may be related to numerous aspects. It is clearly closely related to economy, efficiency, outputs, quality and processes. In addition, it should not be in conflict with objectives related to effectiveness. The productivity phenomenon may therefore play an essential role in improving many of the performance aspects of public services. However, it should also be noted that productivity is not explicitly represented in the models. All of these models may support identifying more detailed success factors based on key objectives of organizations. When paying attention to several aspects it

is possible to achieve a balanced evaluation of performance without too much emphasis on factors such as cost-efficiency. However, the models should not restrict the identification of performance objectives important from the perspective of the organization examined.

The issue that has not been discussed so far relates to cause-effect relationships between different performance factors. Some of the models presented in this section provide information on the assumed relationships between certain aspects such as customer demand and productivity. However, there are also several models, such as the Balanced Scorecard, that do not explicitly present the linkages between different elements of organizational performance. It is clear that many of these relationships are case-specific. Kaplan and Norton (2004) have presented a tool called the strategy map which is a concrete and visual presentation illustrating strategy and reducing the gap between strategy planning and strategy implementation. It presents the assumed relationships between critical success factors and the objectives of an organization.

Knowledge about cause-effect relationships is often related to assumptions based on prior knowledge and experience. However, more statistical research is needed to verify various assumptions, especially in specific services and contexts. There is a need to gain a deeper understanding of the factors affecting productivity (Käpylä et al., 2010). General assumptions on relationships between performance factors are possibly not transferrable to more specific contexts (Bourne et al., 2005). The following figure presents some of the current knowledge on the factors affecting productivity in services. The results of various studies have been classified into the different perspectives of the public service productivity model by Stainer and Stainer (1998).

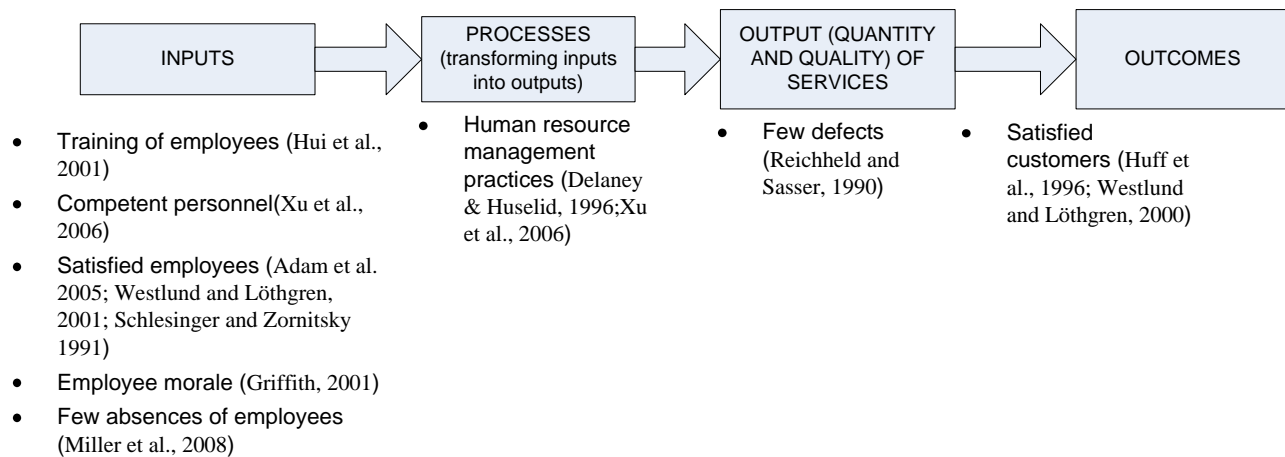


Figure 3 Examples of factors affecting productivity in services.

These studies have concentrated on factors whose role is not self-evident in productivity improvement. For example, the role of quality in the productivity of services is a complex issue with conflicting claims and research results which clearly are affected by the case-specific definition of productivity concept. From these findings it could be concluded that employees are seemingly an important resource for service organizations that need to be managed properly in order to improve their productivity.

1.4 Productivity measurement and management of public services

1.4.1 Phases and challenges in measurement

It has generally been argued that as much as 70 percent of performance measurement initiatives fail (McCunn, 1998). There are several challenges in introducing performance measurement in the public sector. Productivity as a measurement object can cause even more challenges due to problems in actual technical measurement design as well as in committing the employees. In the literature, three phases: design, implementation and use of systems have been identified in the development of performance measurement systems (Bourne, 2003; Lönnqvist, 2004, p. 143, Neely et al., 2000). Designing a performance measurement system entails choosing measurement objects and defining measures, whereas the phase of taking measures into use can be referred to as implementation. Sometimes a separate phase of maintaining (updating and refreshing etc.) the system has also been identified. Even a workable system requires maintaining at times since organizations and their environments constantly change. However, such activities may be related to each of the three other phases (Lönnqvist, 2004, p. 104). Maintaining can include assessment of measurement systems in relation to specified criteria. Criteria for sound measurement have been discussed extensively in the literature (cf. Emory, 1985; Hannula, 1999; Sink, 1985). In this study, four key criteria, namely validity (ability of a measure or a measurement system to measure what it is intended to measure), reliability (consistency of the measurement results, e.g. accuracy and precision), relevance (value and usefulness of the measurement results for the users of the measures) and practicality (cost-effectiveness or the benefit-burden ratio of the measurement), are used to evaluate the soundness of measurement systems developed. In management accounting systems intended for internal use in organizations, relevance is often valued more than objectivity (Kaplan and Atkinson, 1998, p. 1).

Each phase in measurement system development has its own specific characteristics and challenges. This study is related especially to the design and implementation of productivity measures. The first task in designing measures should be related to the identification of purpose of measurement, which is obviously somehow related to the objectives of an organization. The purpose of measurement is naturally also related to the factors to be measured as well as the actual measures. For example, Poister (2003) has noted that public organizations encounter problems in designing performance measurement systems because they are not appropriately designed to serve a particular purpose. Public organizations typically have many stakeholders with different and conflicting requirements (Lawton et al., 2000; Jarrar and Schiuma, 2007). Each stakeholder may have needs related to the purposes and content of measurement systems. There is a risk that taking account of the needs of various stakeholders may result in a measurement system that is too complex and therefore time-consuming to use (Wisniewski and Stewart, 2004).

One underlying cause of problems in the design phase may relate to the definition of a long-term objective. While in companies the ultimate goal is profitability, in public organizations it is often more difficult to define a long-term objective (Rantanen et al., 2007). However, productivity and effectiveness could be regarded as such objectives. In the measurement of productivity different views may be related to the definition of productivity and the role of quality and outcomes in productivity examination. For example, employees and clients of public services may emphasize the

aspects related to service quality and effectiveness whereas taxpayers and top management value high cost-efficiency (Rosen, 1993, p. 1). From the uncertain purpose of productivity measurement (cf. monitoring of productivity trends and identifying means to improve productivity) it also follows that it is difficult to define what to measure.

A specific challenge of productivity measurement in the public sector is to identify a standard output units while every service transaction is sometimes regarded as tailored – customers have a role in the demand for and content of services (Brax, 2007, p. 24, Ojasalo, 1999, p. 59; Rosen, 1993, p. 55; Sahay, 2005). For example, similar services often have different levels of difficulty depending on the customer in question (Rosen, 1993, p. 55). The linkages between outputs and outcomes may also be difficult to identify.

As described in the study by Jääskeläinen and Lönnqvist (2009), the actual design of measures related to the productivity of public services seems to be even more difficult than the identification of measurement objects. Several challenges in the technical design of productivity measures for public services can be identified on the basis of the existing literature. In large public organizations there are often many different departments and units, which means that similar measures and measurement approaches may not apply everywhere (Rosen, 1993). Services may also be inseparable and bundled together in a complex way (Sherwood, 1994). For example, a patient in hospital often receives many services whose role is difficult to capture in a single measure. Outputs of public services are often intangible and qualitative in nature (Flipo, 1988). Intangible and qualitative phenomena are traditionally regarded as difficult to measure (Bourne et al, 2003; Lönnqvist, 2004). In the private sector, output prices have sometimes been proposed as output measures since they also capture the qualitative aspects of service outputs (Grönroos and Ojasalo, 2004). However, this information is rarely available in the public sector (Simpson, 2009). In addition, there are intangible inputs (such as employee competence and information) that can have a role in producing public services. Due to employee-intensity intangible factors may have an essential role in productivity improvement (cf. Xu et al., 2006). However, measuring these factors and taking them as a part of a productivity examination is not easy. Intangible inputs have traditionally been ignored in productivity measurement (Hannula, 1999, p. 39). Finally, a traditional technical problem in productivity measurement is related to commensurating different productivity components (e.g. outputs and inputs).

Potential challenges occurring during the design phase of productivity measurement system development in the public sector are summarized in Table 3. In addition to these challenges, there are also more conventional challenges in productivity measurement, such as the problem of scope (which inputs should be taken into consideration?), allocation (matching principle between inputs and outputs) and accrual (correspondence of output and input allocated to the measurement period) (cf. Hannula, 1999, p. 38). These challenges are not in the focus of this research.

Table 3 Potential challenges in designing productivity measures of public services

Task in measurement system design	Potential challenges
Defining the purpose of measurement	- prioritizing the requirements of different stakeholders (Lawton et al., 2000; Vakkuri and Meklin, 2006; Wisniewski and Stewart, 2004)
Choosing and defining what to measure	- different views on the ultimate organizational goal (also reflected in the different views of productivity definition: efficiency vs. effectiveness) (Rantanen et al., 2007; Rosen, 1993, p. 1) - outputs are difficult to standardize (Gaster, 1996; Ojasalo, 1999, p. 59; Rosen, 1993, p. 55)
Designing of measures	- diversity of services provided by large public organizations (Edwards and Thomas, 2005; Näsi et al., 2001; Rosen, 1993) - inseparability of services (Fletcher and Snee, 1985; McLaughlin and Coffey, 1990; Sherwood, 1994) - intangible nature of outputs (Flipo, 1988; Gupta, 1995; Hannula, 1999, p. 6; Sherwood, 1994) - lack of output prices (Hodgkinson, 1999; Rosen, 1993, p. 4; Simpson, 2009) - intangible nature of inputs (Hannula, 1999, p. 39; Stam, 2007, p. 50) - commensuration of different outputs and inputs (Hannula, 1999, p. 41)

The second phase in measurement development is implementation, which involves the gathering of measurement data, possible test calculations and the development of IT systems. Measurement in itself cannot determine social practices (Vakkuri and Meklin, 2006). It is therefore also essential to inform and train employees and managers in order to gain commitment and ensure the efficient use of measurement systems (Wisniewski and Olafsson 2004). If implementation fails, the potential of measurement system is not realized. In implementing measurement systems, many practical issues have to be determined and documented in order to ensure the successful use of the systems. These include, for example, the purpose of measures, responsible persons related to measures, measurement formulas, frequencies in measurement, target values for measures and reporting of measurement (Neely et al., 1996, p. 64; Uusi-Rauva, 1997, p. 76).

There are several general challenges in implementing measurement systems in the public sector, some of which may be more dominant in applying measurement in specific purposes such as productivity measurement. A general challenge is related to the lack of time and resources for many of the tasks needed. The implementation phase requires a lot of manual work and it is essential to ensure adequate resources for the task (Bourne, 2003). For example, gathering measurement data for questionnaire-based measures takes time. In addition, it has been noted that sometimes there are competing projects squeezing the limited resources in public organizations. Lack of managerial skills has also been related to challenges in implementing measurement systems in the public sector

(Rantanen et al., 2007). In choosing managers, substance skills are often highlighted instead of managerial capabilities. This may have impacts on the training and committing of employees as well as the sharing of tasks related to measurement. There is a risk that the purpose and the actual use of the measurement system are not clearly articulated, which may result in resistance to and mistrust of measurement. This may also be affected by the measurement culture in the public sector, which is considered still to be under development (cf. Ballentine et al., 1998). This is potentially a substantial challenge in productivity measurement, which often has a negative image.

IT systems are important tools facilitating the implementation of measurement systems. They facilitate the gathering of measurement data, carry out calculations and provide reports and visual demonstrations. Problems in these systems are a potential challenge in implementing measurement in the public sector. In addition, it has been found that standards and objectives related to key performance aspects may be lacking in public organizations (Pollanen, 2005). This obviously affects the changes of success in measurement, a key purpose of which is to provide information on the achievement of targets. Lack of standards along with lack of specified responsibilities are two potential causes of challenges in defining practical measurement guidelines and tasks. General challenges in implementing measurement systems in the public sector are summarized in the Table 4.

Table 4 Potential challenges in implementing measurement systems in the public sector

Task in measurement system implementation	Potential challenges
Informing and training employees, Gathering measurement data	<ul style="list-style-type: none"> - time and resources needed (Bourne, 2003, Chan, 2004; Pollanen, 2005) - overlapping projects take resources (Rantanen et al., 2007)
Committing operative employees and managers	<ul style="list-style-type: none"> - resistance to change (Johnsen, 2005; McAdam et al, 2005; Rantanen et al., 2007) - mistrust of measurement, lack of credibility and usefulness (Cavalluzzo and Ittner, 2004; Pollanen, 2005)
Integrating measurement with IT systems	<ul style="list-style-type: none"> - lack of highly-developed IT systems (Chan, 2004; Cavalluzzo and Ittner, 2004)
Sharing tasks and other practical issues related to measurement	<ul style="list-style-type: none"> - lack of standards (Pollanen, 2005) - too many responsible persons leads to non-responsibility (Johnsen, 2005; Rantanen et al., 2007)

The third phase in measurement is the actual managerial use of systems in which measurement results are analyzed and used in decision-making. There may be many challenges in the use of measurement system, many of which result from failures in the previous phases of design and implementation. There may be problems in actual measures, IT systems as well as in sharing of tasks and analysis based on measures. One of the challenges in public performance measurement is

the inefficient use of measurement information (cf. Lancer Julnes and Holzer, 2001). Poister (2003, p. 19) has noted that even if a public organization has a workable measurement system it may not be used efficiently and, at worst, it may even be ignored. Much effort is expended to maintain measurement systems but, at the same time, managers do not necessarily look at the data in a serious way. Similarly, Wisniewski and Olafsson (2004) report that in some cases measurement is regarded as something that must be done with no real understanding of why. These problems may be caused by lack of time and resources or knowledge related to measurement. One reason for inefficient measurement use may also be that the users have no authority to make decisions related to measurement objects (Cavalluzzo and Ittner, 2004).

Even technically sound measurement systems provide unreliable information if there are problems in data collection for the measures (Johnsen, 2005; Wisniewski and Stewart, 2004). Poister (2003, p. 92) has noted that the personnel of public organizations responsible for data input from various offices, branches and work units may end up counting things differently from one another. This may cause problems, especially in large agencies with multiple providers of data. Furthermore, lack of data from single units can result in incomplete measurement results.

Sometimes a problem related to undesired behavior caused by measurement may occur only after using the measurement system for a while. Inappropriate measures or an unbalanced set of measures may result in goal displacement and behaviour that impairs rather than enhances performance (Poister, 2003, p. 20; Vakkuri and Meklin, 2006). In the measurement of public service productivity, the definition of outputs is possibly the most essential task in avoiding this challenge. The output measure of public services should not motivate employees to provide outputs that are easy to achieve at the expense of demanding ones (Rosen, 1993, p. 55).

1.4.2 Managerial and other uses of measurement

Determining the purpose and use of measurement is an essential starting point for the development of any measurement system. It should also be noted that it is often impossible to satisfy multiple purposes with the same measurement system since each purpose represents a specific set of requirements for measurement (Poister, 2003, p. 10; Rosen, 1993, p. 65).

Productivity measurement (and measurement in general) always includes some sort of comparison, since measurement values make no sense as such. The values of productivity measures can be compared to a previous result (trend analysis), results from other similar organizations or units (comparison analysis) or defined target levels (goal analysis) (Hannula, 1999, p. 34; Matta, 1989, p. 66). The main underlying motivation for productivity measurement is obviously productivity improvement. A traditional reason for measuring productivity is to obtain information about the impact of productivity development decisions and measures have been used for control purposes (Hannula, 1999, p. 1, 33). Productivity measurement has also been used in statistical analysis based on data from substantial organizational entities (e.g. Kangasharju et al., 2007). More detailed purposes of productivity measurement seem to be discussed much less. According to Rosen (1993, pp. 58 - 60), the information provided by public service productivity measures may be integrated into decision-making in many ways such as:

- capacity management related to the resources needed for producing certain outputs
- determining efficient ways to allocate staff and other resources
- monitoring the trend of productivity and warning of the emergence of undercutting problems
- demonstrating productivity improvement to the public and politicians
- provision of information related to employee morale, skills, absences and turnover rates for human resources management.

The last of these purposes is obviously difficult to satisfy with traditional output/input measures and is more related to the overall performance of operations. However, the connection between productivity and human resources management as well as rewarding of employees is meaningful and relevant. Since the use of measurement has been discussed in much more detail in the general literature on performance measurement, it is briefly presented next.

In private companies the underlying motivation of measurement is often related to the implementing of strategy by translating it into concrete objectives which can be supported by measures (Kaplan and Norton, 1996, p. 10; Simons, 2000, p. 7; Toivanen, 2001). However, performance measurement is not just a tool for top management. Ukko (2009, p. 11) points out that performance measurement is now focusing increasingly on operative level activities. There may therefore be measurement systems at each organizational level of large organizations. In the context of the public sector, Greiling (2005) has noted that while organizations are not necessarily obligated to develop strategic management systems, performance measurement rarely acts clearly as a strategic steering tool. Greiling (2005), Johnsen et al. (2006) and Poister (2003, p. 10) have identified the following purposes for performance measurement in public organizations:

- modernizing the process of public budgeting and planning (e.g. more focus on output and outcome indicators instead of financial needs)
- monitoring and reporting (e.g. information on the quality of services)
- communication with the public (“report cards” on performance)
- enhancing transparency
- strategic management system (performance measurement as a means to implement strategy)
- performance management (process of directing and controlling employees and work units)
- organizational learning
- contract management (the role of performance measures in service contracts)
- inter-administrative comparison/benchmarking (promoting competition between public agencies).

This list demonstrates the various possibilities for using measurement in public organizations. For example, internal and external use of measurement may be differentiated. There is obviously a case-specific nature in the specific purposes of measurement. The potential of measurement is achieved only if the measurement systems and individual measurement methods satisfy the needs considered important.

1.4.3 Measurement approaches and methods

Many approaches and methods for measuring productivity have been presented in the literature. Much of the discussion on productivity measurement has been carried out by economists with a specific interest in the productivity of a national economy (Phelps, 2009; Simpson, 2009). In addition, the topic has been examined in a manufacturing context (Hannula, 1999; Saari, 2000). It is difficult to establish a coherent and comprehensive picture of productivity measurement since it has been developed in different academic fields (economics, management, accounting etc.) and for various purposes. For example, different measures are required depending on the level of examination (cf. national economy and an individual department of an organization). The focus of this study is on measurement supporting productivity management. It is examined first by presenting general approaches. Then the examination is focused on index methods and multidimensional measurement. Macro level productivity measurement and productivity analysis methods are briefly reviewed at the end of this section.

General measurement approaches

Basically with every productivity measure there are two alternatives: measuring the level of or the change in productivity. These can also be referred to as static and dynamic productivity measures (Sink, 1983). Static measures (e.g. number of customers served per number of employees) can be used, e.g. in benchmarking the productivity level of similar organizational units at a certain point in time. Dynamic measures compare the present result of the measure to a former result of the same measure. The results of dynamic measures are typically easier to compare between different activities. Mammone (1980) presents two approaches to the productivity measurement of organizations: component (disaggregated) measurement and aggregated measurement. Component measures capture the productivity of a single activity or unit. Aggregate measures examine the productivity of a large organizational entity or the whole organization by examining multiple inputs and outputs simultaneously. One example of a method applying an aggregated approach that has often been used in the public sector is Laspeyers index, which first aggregates different quantitative outputs and inputs and then examines the ratio between them (Simpson, 2009).

Somewhat similarly to Mammone (1980), McLaughlin and Coffey (1990) suggest that the measurement of service productivity could be first focused on the disaggregated components (e.g. product or process) of a complex mix of services. In this approach, inputs and outputs for each component are examined separately and compared separately. McLaughlin and Coffey also argue that much of the emphasis needs to focus on disaggregate measures due the complexity of productivity in services. Kaydos (1999, p. 85) also suggests that it is meaningful to concentrate on separate processes or functions in productivity measurement since it is difficult to capture all the outputs and inputs of large organizational entities. More generally, Kald and Nilsson (2000), among many others, have stressed the need to extend measurement organizationally downward, since the relevance of measures increases, which may improve the acceptance of measures.

However, component measurement of productivity is often not enough since information on the productivity of the whole organization is needed (Simpson, 2009). Therefore, the results from component measures should still be somehow aggregated. This causes a commensurability problem:

how to sum up the results from various component measures. The commensurability problem of productivity measurement has traditionally been related to summing up various inputs (cf. material and labor) and outputs (e.g. different products). In practice, the only feasible way to commensurate various inputs and outputs is to use the monetary values of the inputs or outputs in question (Craig and Harris, 1973).

In the context of industrial production, partial productivity measures have often been used. Hannula (1999, 2002) has proposed a method which aggregates the results from different partial productivity measures (“productivity components”) resulting in the measurement of total productivity. The basic idea is first to measure the change in partial productivity ratios. Then the results are aggregated by using information on the costs and prices of individual productivity components. Productivity change is examined in order to keep the results from different measures comparable.

Index measurement

Various indices are the most common way of measuring productivity (Singh et al., 2000). They concentrate on the examination of the ratio between output and input and may also be called direct productivity measures due to equivalence with the definition of productivity. Many variations of the output/input index may be applied, such as total productivity (all outputs / all inputs), partial productivity (all outputs / one input, such as labor) and physical productivity (essential output / essential input) (Craig and Harris, 1973; Uusi-Rauva, 1997, pp. 44 - 56). In addition to the commensuration of various inputs and outputs, there are two general issues in index measurement: how to measure inputs and outputs as well as how to allocate inputs to outputs. The former is discussed next in more detail.

Traditionally inputs have been related to labor (L), capital (C), material (M) and energy (E) and measured respectively with person-hours, person-years (L); monetary values (e.g. depreciation) related to assets on the balance sheet such as machinery and premises (C); units of material inputs (e.g. kilograms, liters) (M) and watt-hours (E) (Craig and Harris, 1973; Hannula, 1999, p. 62; Uusi-Rauva, 1997, pp. 49 - 54). In addition, intangible inputs such as information have sometimes been considered. For example, the skills of employees clearly cause variation in the labor inputs used. However, it is hard to find any methods that specifically capture intangible inputs in productivity measurement. In the public sector, the measurement of inputs is not usually regarded as a challenge since costs are used to measure inputs used (cf. Faucett and Kleiner, 1994; Hodgkinson, 1999). The problem in this approach is that costs may be affected by factors external to productivity and beyond the control of an organization. However, using monetary values for measuring inputs is often practical since the existing cost accounting can be utilized as such. For example, unit costs (costs / outputs) can provide relevant information on productivity when fixed value for money is used (Stainer and Stainer, 1998; Uusi-Rauva, 1997, p. 100).

As discussed earlier, capturing outputs is a key challenge in the productivity measurement of public services. The eventual validity of public service productivity measurement is essentially related to the outputs used in measurement (Häkkinen, 2008, p. 110). In general, four ways of measuring outputs may be considered (Häkkinen 2008, p. 97; Grönroos ja Ojasalo, 2004; Sherwood, 1994):

- quantity of service transactions

- quantity and quality of service transactions
- outcomes
- monetary measures.

In practice, quantitative outputs related to service transactions such as care days are commonly used in the measurement of public service productivity. They are direct or immediate outputs of services (Gadrey, 1988). These outputs typically do not capture aspects related to the quality of outputs. Differences in customers are not taken into account, meaning that measurement may be called “non case-mix adjusted” measurement (Häkkinen, 2008, p. 97). Due to the practicality of this approach, it may be suitable if quality is assumed to be constant in all the units providing similar services or does not change over time (Rosen, 1993, p. 91, 102 - 103).

There are various ways of capturing different aspects of service quality. First, different classifications can be applied with separate classes for output transactions of different difficulty (cf. technical quality by Grönroos, 2001). According to Rosen (1993, p. 91), similar outputs of different difficulty can be counted together by first taking the least difficult unit of work as a standard unit of work. After that the other more demanding classes are weighted in relation to the time and effort needed to produce such outputs. Studies on the productivity analysis of health and elderly care commonly use such customer-structure adjusted outputs with specific methods such as RUG (Resource Utilization Groups) and DRG (Diagnosis Related Group) (Häkkinen, 2008, pp. 97–98, 110). In these sectors, quality changes may be so rapid that weights related to customers have to be updated every two or three years. There are positive experiences of using such methods in improving the financial control of operations (e.g. Modell, 2001). Second, quality measures can be used in excluding or discounting quantitative outputs that fall below a desired quality standard (Rosen, 1993, pp. 102 - 103). In practice, the discounting may be carried out by comparing the realized value of a certain quality measure to a defined standard value. For example, an average of 3.5 in a customer satisfaction survey per standard of 4.0 = 0.875. Quantitative outputs are multiplied by the value obtained. These ideas have potential in principle but the practical application is not easy due to lack of quality information and the effort needed for using such measures.

The measurement of outcomes has also been proposed as a solution to the problems in measuring service outputs (e.g. Gadrey, 1988; Sherwood, 1994). The intention is to measure the effects (e.g. learning) or the results of service transactions (e.g. educational course). In this approach, it is not necessary to identify all the factors related to service transactions. Only a limited number of measurable benefits is required. In addition, this approach makes it possible to capture certain quality-related factors without any specific measures. Outcomes of public services may occur instantly after service transaction or after an extended period of time. As Häkkinen (2008, p. 97) describes it: solving the problem of a patient is often affected by various hospital services resulting in a healthy patient. On the other hand, longer term health impacts may be related to extended years of patient life. The latter types of outcomes are especially problematic from the perspective of productivity measurement. It is difficult to distinguish the influence of service provider from other external factors. Health in the long term may be affected by many decisions made by the patient.

Many of the traditional productivity measurement methods (e.g. Davis and value added methods) require monetary values related to outputs (see e.g. Saari, 2000). In the discussion on service productivity, monetary output measures have also been proposed (Grönroos and Ojasalo, 2004; Klassen et al., 1998; Viitamo, 2007). This has been rationalized by stating that the quality of services and the value from the perspective of customer can be better captured. Only payment received from a service is considered to contribute to the output. Even though practical and potentially useful, this monetary approach should be used with care in order to be able to distinguish productivity and profitability. A further obstacle is that detailed price information may not be available, which is a problem especially in the public sector.

A broad approach

It is difficult to design a single index comprehensively capturing all the aspects related to the productivity of public services. Even if this is achieved, there is a risk that the result is complex and difficult to use in daily management. In this study, the term broad approach is used to describe measurement supporting productivity management but not necessarily directly examining productivity. It may include surrogate or indirect productivity measures which are related to factors that are highly correlated with productivity but do not represent the ratio between output and input (cf. Sink, 1983). Indirect measures may be related to factors reflecting both high (employee satisfaction) and low productivity (absenteeism).

Productivity measures (e.g. direct, indirect, partial and component) may serve as a part of balanced performance measurement systems based on frameworks such as the Balanced Scorecard, Performance Pyramid or Performance Prism (cf. Kaplan and Norton, 1992; Lynch and Cross, 1995; Neely and Adams, 2000). Other related factors such as effectiveness, processes, demand for services as well as motivation and competence of employees may be measured with separate measures (cf. Rosen, 1993, p. 58, 102 - 103). A motivation for multidimensional measurement is that it enables the gathering of information relevant in identifying concrete means for improving productivity. It can also be rationalized with the complexity of public service productivity as a measurement object (Gupta, 1995). A problem in this approach may be that productivity-related factors remain disconnected in practice. In addition, the core of productivity may be forgotten or ignored in designing performance measurement systems (Hodgkinson, 1999).

A broad approach may also be applied in a separate measurement system for productivity. This can be carried out by using a measurement method called a productivity matrix¹ (Riggs, 1986) which has been successfully applied in many industries (Dervitsiotis, 1995; Rantanen and Holtari, 1999). The productivity matrix (see an illustration in Appendix 1) is a method in which a set of direct and indirect (e.g. number of quality errors and employee sickness absences) productivity measures are used to compose a single measurement score. Every measure in the productivity matrix has its own weight (0-100) in the calculations. In a traditional application of the matrix, the expected values of different measures are scaled in order to produce a score from 0 to 10 for each measure (Allender,

¹ Sometimes also referred to as the objectives matrix (OMAX), multi-criteria performance measurement technique (MCP/PMT) and importance-performance matrix

1997). By first multiplying the score of each measure by the weights and then summing up the results, the matrix produces a total score from 0 to 1000. The matrix method differs from the conventional productivity examination and as such may be regarded as an indirect productivity measure.

Macro level and analytical methods

There are also many other ways of productivity measurement in the public sector. These are related to the measurement of productivity at the macro level as well as various methods for analyzing productivity. Statistics Finland provides public sector productivity change figures for state and local administration (educational, health and social services) (Hautakangas et al., 2007). Total output of local administration was earlier examined solely with deflated costs, meaning that these services were considered as non-value adding entities. This has also been the case in many other countries. However, in more recent calculations in the Finnish public sector an examination of change indices related to quantitative outputs has been applied in non-collective services. (Boyle, 2006; Statistics Finland, 2006) Laspeyres index is used to produce a change index for output and input:

$$Q_t^{t+1}(La) = \frac{\sum_{i=1}^n p_i^t q_i^{t+1}}{\sum_{i=1}^n p_i^t q_i^t}, \text{ in which}$$

p_i^t = weight of product i in base year t (based on proportion of costs)

q_i^t = quantity of product i in base year t (output: e.g. number of customers served; input: costs incurred deflated to base year value for money)

The method used in the state administration is similar to that used in local administration with a few exceptions: collective services (no direct customers, e.g. national defence) are also examined, productivity of labor is also calculated, the Divisia-Törnqvist-index is used to produce change indices for output and input. The calculations provided by Statistics Finland do not provide information on the effectiveness of public services or the necessity of various services. (Hautakangas et al., 2007)

In the analysis of public service productivity, methods such as data envelopment analysis (DEA) and stochastic frontier analysis (SFA) have been applied (Hollingsworth, 2008; Simpson, 2009). These methods are typically used for examining differences in productivity between organizations or other organizational entities. The efficiency term is used in examining these differences (Kangasharju et al., 2007, p. 126). These methods are more often used in the examination of the productivity level instead of change. It seems that DEA especially has been widely utilized in public sector productivity studies (see e.g. Jacobs, 2001; Ruggiero, 2006; Vakkuri, 2003). According to a study by Hollingsworth (2003), 50 percent of 188 efficiency measurement studies examined in the context of health care applied DEA analysis whereas 12 percent applied SFA.

DEA analysis is suitable for comparing the productivity of similar organizations (Simpson, 2006; Vakkuri, 2003). In DEA there is no need to understand the formula between outputs and inputs.

DEA has proved its value as a benchmarking tool between similar organizations (Uusi-Rauva, 1997, p. 74). DEA is a flexible, data driven and non-parametric linear programming technique in which the production frontier is constructed on the basis of different combinations of output/input ratios. The contribution of various inputs (e.g. computers and teachers per pupil in schools) is evaluated in relation to the achievement of a given set of outputs (e.g. pupils' test scores in different subjects in schools). (Singh et al., 2000; Vakkuri, 2003) The production frontier represents the maximum output that can be provided with the given input. Inefficient organizations are found to be below this production frontier. (Uusi-Rauva, 1997, p. 74) Efficiency of organizations or units is compared to a 'fully' efficient counterpart (optimal output combination with the given set of inputs) (Simpson, 2009). In practice, the percentage of the inputs needed if an organization was as efficient as the most efficient counterpart (or counterparts) can be clarified. Similarly, it is possible to investigate how many percent more outputs could be produced with the inputs available if the efficiency was the same as in the most efficient counterpart.

There are several advantages with DEA, such as the possibility to use an extended set of inputs and outputs (e.g. there can be several output criteria at the same time, not just one), no need for a specific functional form, causality or error term, as well as no need for monetary data related to inputs and outputs (quantities can be utilized as such without weights) (Simpson, 2009; Singh et al., 2000). If there are many outputs and inputs it is likely that more organizations will be at the production frontier due to some specific variables. When many variables are used it is more likely that there will be a unique output/input combination indicating efficiency in the method. If there are no similar organizations in the data set, it is possible that an inefficient organization is at the production frontier. If there are many variables, there is therefore a need for a large number of organizations in the data set. (Kangasharju et al., 2007; Niemi, 1997) Hence, there are at least two disadvantages with the DEA method: the requirement of a large series of data and difficulty in developing tests of statistical significance due to the non-stochastic procedure (Niemi, 1997; Singh et al., 2000).

SFA is a parametric estimation technique which, like DEA, models the production frontier. It has its roots in regression analysis and can be used to measure the inefficiency of each examined organization relative to the estimated 'optimum'. (Simpson, 2009) There is a need to understand the production system in order to estimate the production frontier, which is not a straightforward task. With SFA it is easier than with DEA to take into account various external factors (e.g. socio-economic factors) which organizations cannot affect. In addition, the impact of random variation (extremes in the data set) can be separated from efficiency examination and there may be no efficient organizations at all. (Kangasharju et al., 2007)

An essential challenge in any method concerns the outputs used. The quality of data used in the macro level calculations of public sector productivity has often been criticized. If inaccurate data is used, the measurement yields invalid information, which may lead to erroneous conclusions. The productivity analysis methods such as DEA are more oriented towards the purposes of research and analysis (e.g. non-recurring study of factors affecting productivity) than managerial use (cf. Niemi, 1997). They require large data sets from many organizations or other organizational entities, which limits their potential for use on the operative level.

2 RESEARCH DESIGN

2.1 Research gap, objective and questions

The existing research including many known challenges related to productivity measurement and management in public service organizations was already discussed in detail in the first chapter of this study. The purpose of this section is to first describe the research gap to which this research aims to contribute. At the same time, some key research topics underlying this study are briefly described. Finally, the research objective and detailed research questions are presented.

An underlying motivation of this research is to provide further understanding of measuring productivity in public service organizations. However, the measurement object must be understood before the actual measurement. In the discussion of productivity it should be noted that the concept is operationalized very differently. The practical meaning of the concept seems to depend on the context within which it is used (Linna et al., 2010; Tangen, 2005). *Productivity research* is very diverse (Käpylä et al., 2010). A challenge of the multidisciplinary field is scattered knowledge and absence of co-operation (Rosen, 1993, p. 4). Pritchard (1995, p. 2) has presented a classification which aptly describes the various aspects of productivity research:

1. Economist/engineer approach, in which productivity is seen as an efficiency measure (outputs/inputs)
2. Approach in which efficiency (outputs/inputs) and effectiveness (outputs/goals) are evaluated simultaneously
3. Broad approach, which comprises everything that makes an organization function better

Productivity may also be examined at various levels. Basically two main levels can be identified: a) the macro level examination relates to the international, national economy or industry level; b) the micro level perspective can be related to a single organization, department, unit, process or individual employee. The background of the examiner (e.g. politician, economist, accountant, behavioral scientist, engineer, manager) is related to the productivity aspect used as well as the level of examination. (cf. Hannula, 1999)

Issues related to the *productivity phenomenon in services* have been investigated in various studies which examine the topic especially from the perspective of the various specific characteristics of services (Grönroos and Ojasalo, 2004; Parasuraman, 2002). It has been stated that there is a lack of empirical research on the topic (Gummesson, 1998; Vuorinen et al, 1998) focusing on micro level examination (Brax, 2007, p. 2; Linna et al., 2010). In addition, many of the existing studies have focused on the private sector. This study aims to contribute to the existing discussion with the context of the public sector and with the empirical emphasis on micro level operations.

Management accounting systems assist several key roles within organizations by enhancing decision-making, guiding strategy development, evaluating existing strategies as well as focusing efforts related to evaluating and improving performance. Examples of activities carried out by management accounting are collecting, classifying, processing, analyzing and reporting information

to managers. (Kaplan and Atkinson, 1998, p. 1, 12) One of the topics which has rapidly attracted more interest in the field of management accounting relates to *performance measurement*. Much of the general knowledge related to measurement utilized in this study is based on performance measurement research. This research has provided companies and public organizations with tools for designing measures that satisfy managerial needs. Balanced performance measurement systems include both financial and non-financial measures related to factors important from the perspective of organizational performance (Kaplan and Norton, 1992; Lynch and Cross, 1995; Neely and Adams, 2000).

Performance measurement and management in the public sector is a topic that has been studied a lot during the last two decades. According to Van Helden (2005), 37 per cent of the studies on public sector management accounting examine performance measurement. One reason for the multidimensional performance evaluation of public services relates to the fact that improvements in some areas (e.g. cost efficiency) may simultaneously impair other areas (e.g. service quality or ultimate outcomes) (Boyne, 2003). There is both theoretical and empirical research on the design (Edwards and Thomas, 2005; McAdam et al., 2005; Wisniewski and Olafsson, 2004), implementation (Cavalluzzo and Ittner, 2004; Collier, 2006) and use of measurement systems (Greiling, 2005; Chan, 2004; Vakkuri and Meklin, 2006) as well as various frameworks of organizational performance (Boyne, 2002; Talbot, 1999). As with the research on productivity, these issues are studied in different disciplines (e.g. economics, management accounting, public administration, sociology) with limited co-operation (Van Helden et al., 2008). The studies on this topic comprehensively examine various dimensions of public service performance without specific emphasis on productivity. Several challenges have been identified in the development of measurement systems in the public sector (see e.g. Chan, 2004; Poister and Streib, 1999; Rantanen et al., 2007). These challenges are caused by many factors, such as diverse goals of stakeholders and may also be dependent on contextual factors. It may be argued that some of the challenges in developing productivity measurement systems are even more severe due to resistance to change and the technical issues of measurement. There is a lack of research examining measurement system development in such a specific setting.

Despite the importance of productivity, issues related to the *measurement of productivity* have still not received adequate attention (Singh et al., 2000). Singh et al. (2000) examined 23 studies on productivity measurement, the majority of which were carried out in the manufacturing context and had a statistical and analytical approach focusing on industry or firm level examination. In 1989 Banker et al. noted that much of the research on productivity measurement had been carried out by economists or industrial engineers and there were only a few studies on the topic in the field of management accounting. This still seems to be the case in the 21st century. There are many issues related to the topic waiting to be solved. Many challenges have been related to measuring productivity of services and it has been argued that the understanding of service productivity measurement is still in its infancy (Grönroos and Ojasalo, 2004; Gupta, 1995). This has been explained by the specific nature of services, meaning that traditional ways of measuring productivity would not be appropriate. It has also been stated that service organizations have not applied productivity measures in a similar manner in comparison to manufacturing companies (Van

Looy et al., 1998). In addition, the connection between productivity measurement and management seems has attracted surprisingly little attention.

Many of the existing studies on the topic of *public service productivity measurement* are theoretical (Hodgkinson, 1999; Larsson, 1995) or take an analytical approach focusing on larger organizational entities (Lorenzo and Sanchez, 2007; Phelps, 2009; Simpson, 2009). These studies support political decision-making and macro analysis but are of limited use for managerial purposes at the operative level. The need for more detailed and accurate basic information has also been voiced in the discussion on macro level productivity (Maliranta, 2007, p. 45; Ministry of Finance, 2007). The health care sector has also attracted a specific attention in operative level examination (see e.g. Gupta, 1995; Hupli et al., 2006; Kujala et al., 2006; Nordgren, 2009). One reason may be the similarity of some of the operations with manufacturing processes. Even with this sector, it has been stated that the understanding of productivity-related issues is underdeveloped (Berry and Bendapudi, 2007). This study aims to contribute to the discussion on public service productivity measurement especially from the managerial perspective, meaning that the interest is not only in the productivity figures (output/input) as such but also in the underlying factors (e.g. why productivity is good or poor?). There is a lack of systematic examination of the various alternatives to utilizing measurement information in productivity management.

Many of the factors affecting public service productivity may be intangible. Producing services is employee-intensive, which means that factors such as employee competence and working atmosphere have a role in inputs. On the other hand, output of services may include intangible elements, which has been considered to be a fundamental reason for productivity measurement challenges (Flipo, 1988; McLaughlin and Coffey, 1990; Sherwood, 1994). Challenges in measuring and managing intangibles have created a new *intellectual capital* (IC) research field which has produced a lot of literature on identifying, measuring and managing intangibles related to individuals, organizations and societies (see e.g. Carson et al. 2004; Chatzkel, 2004; Lönnqvist, 2004; Kujansivu 2008). The discussion on intangibles seems to focus on intangible resources or inputs (Carson et al. 2004; Edvinsson and Malone, 1997; Sveiby, 1997) whereas intangible outputs or outcomes have been studied less (Jääskeläinen and Lönnqvist, 2008). This study aims to contribute to the discussion on capturing the intangible elements of service outputs. It also applies existing ideas proposed for measuring intangibles in the context of public service productivity measurement.

To summarize, there are various practical challenges and research needs regarding the topic of public service productivity measurement and management. The aim of this research is *to investigate the development of measurement systems supporting the productivity management of large public organizations*. The structure of this study draws loosely on the commonly known phasing of the development of performance measurement systems which includes design, implementation and use of measurement systems (Bourne, 2003; Lönnqvist, 2004, p. 143; Neely et al., 2000). First, productivity phenomenon must be understood in order to later identify concrete measurement objects. Second, the needs for measurement must be determined to enable relevant measurement objects to be defined and selected. Third, measures can be designed after which the measurement system is ready to be tested and implemented in managerial use. Finally, the whole development

process may be evaluated. Due to the practical limitations, the phase related to the actual use of measures is not within the scope of this research. Since public service productivity itself is an ambiguous phenomenon, it is extensively studied. Consequently, this research focuses on two main research questions which are divided into subquestions:

1. How can public service productivity be described as a phenomenon?
 - a) What affects productivity at different organizational levels?
 - b) What is the impact of different productivity factors at the operative level?
2. How can productivity management be supported by means of measurement?
 - a) What kinds of needs for measurement can be identified at different organizational levels?
 - b) How can the needs be satisfied with measures?
 - c) What affects success in the process of developing measurement systems?

Productivity factors include various elements, such as employee satisfaction or output quality, that may play a role in improving the productivity of public services and may therefore be regarded as essential from the perspective of productivity management. Measurement needs are related to various purposes (e.g. benchmarking of units) and criteria (e.g. practicality) for measurement.

2.2 Scope of the research

As discussed earlier, productivity research is very multidisciplinary. This study is carried out in the field of applied business economics and industrial management. The research issue is therefore examined from the perspective of management. Consequently, factors external to the power of public managers such as the results from political decision making (e.g. wide perspective towards structures of organizing services), are not of primary concern. The productivity concept is seen from the traditional engineering perspective (as a relationship between output and input) with the emphasis on micro level examination. The service provider's perspective is used in productivity examination. In line with the literature on (public) service productivity, however, output quality aspects are highlighted. When operationalizing the productivity concept for purposes of measurement and management, it is also seen from a broader perspective since there are many factors that may have a role in improving productivity.

The managerial perspective of this study means that the internal use of measurement is investigated rather than external reporting. Productivity and performance measurement has often been linked to the rewarding of employees. Since rewarding is a wide topic, it is not specifically addressed in this study. The study focuses on designing and implementing measurement systems. For practical reasons the actual use or impacts of measurement are not within the scope of this research. Meaningful examination of these issues could be possible only after years of experience using the systems developed.

The aim is to keep the research as generalizable as possible. However, the case setting inevitably affects the scope of this research. The empirical context of this study is a large municipal organization providing and arranging a wide range of public services. This means that the results of the study can be of only limited use in small organizations (fewer than 50 employees) providing

only few services. Productivity-related issues can be examined at various organizational levels of large public organizations. In this study, issues are simplified by examining only top (municipal organization as a whole or municipal department providing many different services) and bottom (administrative unit providing few specific services) levels. There are different ways to affect productivity at these levels. At the top level it may be examined which service forms are emphasized in striving to achieve mission as well as the ways of organizing certain services. (e.g. where they are provided and by whom) (cf. Laine, 2005). The top level examination of this study is limited to two research articles.

The ways to affect productivity at the bottom (operative) level may be related to how catering and property maintenance are organized, how work is distributed in an optimal way and how existing capacity is adjusted to changes in demand (ibid, 2005). The main emphasis of this research is on the operative level, on which various services are produced. The operative level examination is carried out in the context of social services. The most detailed examination therefore focuses more on welfare services, which means that the application of the results to some other public services, such as technical services, may not be straightforward. Even though the operative level units such as child day care centers may be small as such, the very fact that they are a part of a larger organization inevitably has an impact on the scope of this study.

Services have been privatized and the purchaser-provider model has been implemented in many Finnish public organizations. However, discussion related to these topics is not in the specific interest of this study due to the limitations of the empirical examination. Finally, the characteristics of the Finnish public sector, such as the legislation and the arrangement of services have some effect on the results of this research. The general characteristics of management in Nordic countries are egalitarianism, hearing the opinions of all the stakeholders and incremental change instead of radical reforms (Johnsen et al., 2006). The Finnish public sector is similar to that in Scandinavian countries characterized by democratic decision-making and a welfare state (Pollitt and Summa, 1997). There seems to be no specific reason, however, why many of the results could not be applied in the public organizations of other western countries if these characteristics are acknowledged.

2.3 Description of the empirical context

The aim of this section is to briefly describe the empirical context of this study, the main aims of the practical development work and the history of productivity measurement in the organization studied. This study was carried out along with two research and development projects during 2007-2010 in which the author acted as the main researcher and project manager. The projects included development work in productivity measurement in the City of Helsinki, Finland. The starting point of the work was deficiencies identified in the measurement in use. Representatives of the central administration of the City of Helsinki asked the Performance Management Team research group at Tampere University of Technology to provide external expertise in the development work. This setting enabled various research activities to be linked to the practical development work. The practically-oriented research project typical in the research tradition of industrial management had a perceptible effect on the action research approach chosen for this dissertation. In this way it was possible to gain optimal access to relevant empirical data (cf. Andriessen, 2004) and management

reality (cf. Gummeson, 2000, p. 25) and it was also appropriate from the perspective of the existing literature.

The City of Helsinki is a large municipal organization with around 40,000 employees and an annual expenditure of around €3,000 million. Social welfare and health care account for more than a half of the total expenditure. Most of the income is derived from tax revenues. There are around 40 departments such as City Transport, the Cultural Office, the Education Department, the Health Centre and the Social Services Department. The City Council is the supreme decision-making body of the organizations and delegates power to the City Government, various committees and individual officials. (City of Helsinki, 2002) The structure of the organization is described in Figure 4.

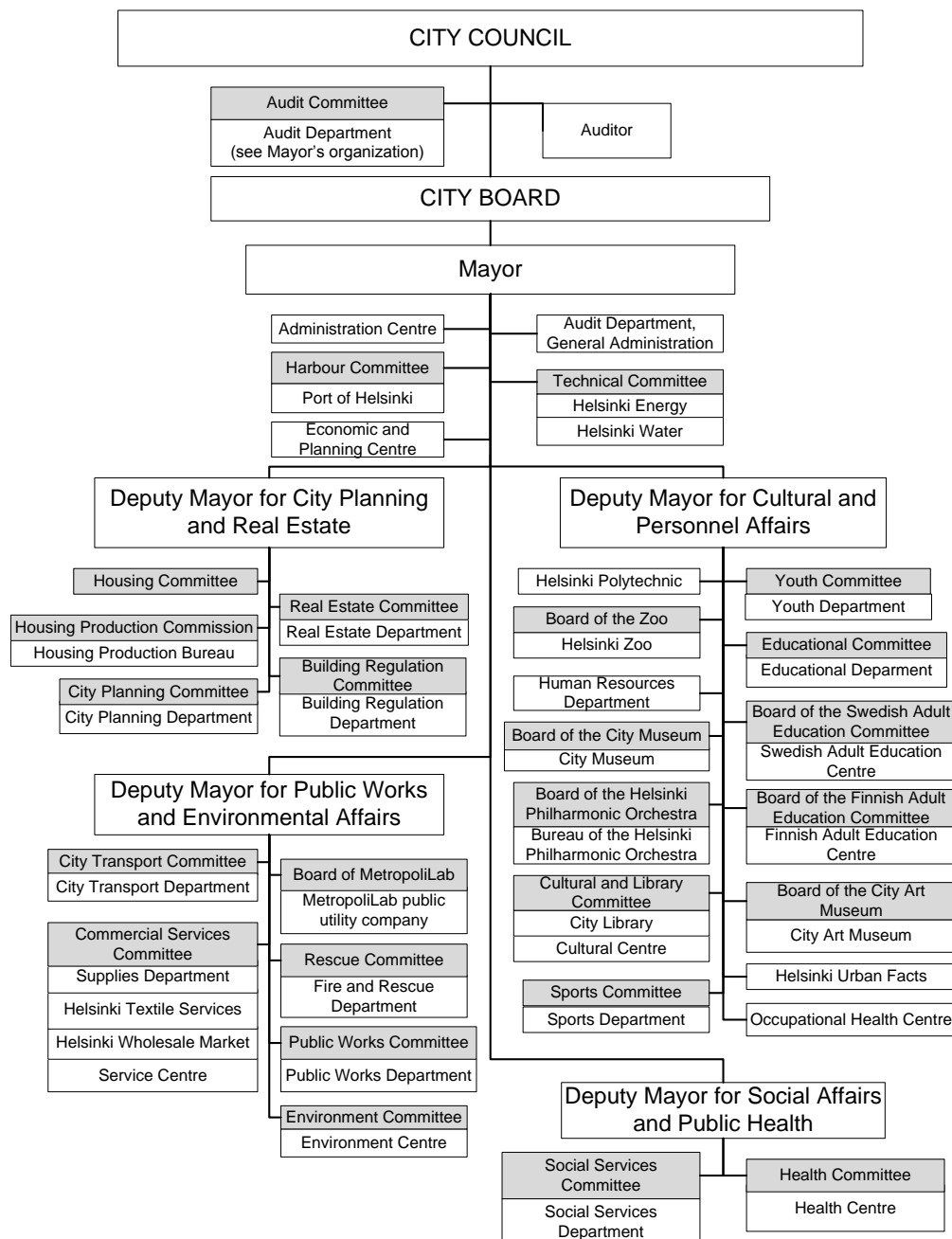


Figure 4 Organizational structure of the City of Helsinki (City of Helsinki, 2008)

The city organization as a whole includes a huge variety of operations and could be described as a multi-branch concern with no equivalents in the private sector. Common denominators of the operations are related to the provision of services instead of products likewise limited market mechanism and lack of proper prices for many of the individual services. In addition, most of the services strive to accomplish a mission such as improving the welfare of residents instead of the typical goal of satisfying shareholders in companies. Public corporations (part of the city concern) providing services such as transportation and energy are special cases which are naturally closer to the private sector organizations.

Prior the research activities described in this research, productivity measurement had been conducted for years in the City of Helsinki. Since various documents and publications provided only a superficial picture of the history of productivity measurement, a productivity measurement expert working in the central administration of the City of Helsinki was interviewed in spring 2009. The interviewee was a key figure in developing the second “wave” productivity measurement (the system used before this research and development work was initiated) in the organization and had 35 years’ working expertise in the City of Helsinki. A single interview lasting for one and half hours was regarded as reliable enough for the purposes of improving the information content of written documentation (cf. Voss et al., 2002). During the transcription of the interview data, specific issues related to schedules and technical measurement details were checked against written documents and corrected where appropriate. The main results of the interview are briefly presented next.

The first “wave” in productivity measurement was in the late 1980s. An external consultant company was hired to develop a system for an analysis of productivity trends. Inputs were measured by costs and outputs with various means such as number of service transactions (number of customer visits etc.) and volume of service provision (number of care places etc.). In some cases, various ratios, such as the number of personnel (or costs) of a specific department in proportion to the number of personnel in the whole city organization or population in the city region, were used. The bottom level for examination was departmental level. The actual results of the measurement based on historical data (1980-87) generally suggested a decreasing trend in productivity. According to the interviewee, representatives of departments had no role in the development of the measurement system used but they were given an opportunity to comment the results of the calculations.

In the first wave, the scope of measurement was at a good level in all departments. According to the interviewee, the key benefits of the first wave could be related to the start of discussions on productivity. Key concepts related to productivity and the difference between effectiveness and productivity became clearer. However, no actual targets for productivity development were identified. None of the departments continued productivity measurement with the methods used. Productivity measurement was met with fierce resistance and commitment towards it was never achieved.

At the beginning of the 90s, the second wave of productivity measurement was initiated. One key driving force was a productivity campaign that started at a time of economic depression in Finland. According to the campaign, the productivity of whole organization should be improved for 10 percent within three years time. In order to monitor the achievement of the objective, productivity measures were needed. It was also considered that measurement by results, which had just been launched, required other 'hard' measures than costs. The measurement method used was mainly developed by the interviewee, a productivity expert working in the central administration. Accounting experts from individual departments were also communicated from early on. There was a separate expert in the city central administration who developed the measurement used in public utility companies.

The use of productivity measurement was made mandatory by an official resolution. Experts in individual departments applied the measurement method and developed their systems based on the instructions provided. The bottom level for examination was mainly departmental level. Productivity change was calculated on the basis of information on the ratio between outputs (aggregated result of various quantitative outputs such as number of customer interactions) and inputs (measured with total costs). Outputs were defined in a more detailed and systematic way than in the first wave. It was instructed that outputs should be related to services provided for the public – not internally between or within city departments. This was in line with the customer perspective of management by results. The productivity data from departments was aggregated in order to reach the level of the whole city organization. Around 90 percent of the operations were covered with the measurement results. Neither central nor departmental administrations were included in the calculations. According to the results, productivity was generally decreasing with a few exceptions. The decreasing trend was not as dramatic as in the first wave calculations.

The measurement survived the period of the productivity campaign well. The directors regularly investigated productivity-related issues during the three-year period. However, later in the 90s the use of productivity measurement faded and it never achieved a full commitment from the field. Between 1991 and 2000 city-level figures were calculated. At the end of that period, the evaluation of the status of measurement revealed certain unclarities in the results. Thereafter the departments measured their productivity in different ways and the results were not calculated at the city level. Departmental budgets included productivity figures but they were generally not used in a systematic manner.

According to the interviewee, a problem in the second wave measurement may have been that it was not properly applied in the departments. Departmental measures were often inadequate. Another problem was that measures did not properly capture aspects of service quality. Another key problem was perhaps that productivity measurement results were not analyzed systematically (e.g. the reasons for certain productivity changes). They were often examined by controllers and other accounting experts but not by managers. Often no official productivity targets were set either, which may be the reason why measurement did not function properly as a part of management by objectives. The interviewee described the status of productivity measurement before the development work started in 2007 as a 'calculatory curiosity'. Only in specific contexts such as public utility companies were measures really in managerial use. As a final comment outside the

interview, it should be noted that despite the challenges described above, the City of Helsinki had a long history of productivity measurement which could still be described as advanced (prior to the start of the development work in 2007) compared to many other similar organizations in Finland (cf. Linna et al., 2010).

The status of performance measurement in general is also interesting from the perspective of productivity. Balanced Scorecards were introduced in Helsinki in 1995 along with payment by results. According to the interviewee, however, individual departments or services were not obligated to develop and use performance measurement systems. Consequently at the time of the interview they were not used in all the operations. Payment by results was used in around one third and balanced performance measurement in around half of the whole city organization. These systems were also implemented on the lower organizational levels of large departments as opposed to the departmental level mainly used in productivity measurement.

During an informal discussion after the interview, an issue related to the culture of comparisons was discussed. According to the interviewee, comparisons between departments or units were quite rare in the organization in the early 1990s. In practice, the public presentation of productivity change indices from different departments was a daring move. At that time, not even similar services were compared to each other. At the time of the interview, the interviewee considered that the organization was more ready for comparisons. This trend has also been described in the literature (e.g. Ammons et al., 2001).

In the research and development project started in 2007 the Social Services Department (see Figure 5) was chosen as a specific context for productivity measurement development work. Due to the lack of resources it was not possible to comprehensively examine several departments. However, the perspective of the whole city organization was kept constantly in mind. The Social Services Department was chosen due to its significance from the point of view of productivity – the department is clearly the largest in Helsinki (and also in Finland) and consumes an essential part of the financial resources. There are 12,000 employees in the Social Services Department with an annual expenditure of €1,000 million (Sosiaalivirasto, 2008). In addition, it was regarded that there were most challenges in the measurement of productivity in social services and if these challenges could be solved the new knowledge could be applied in other departments.

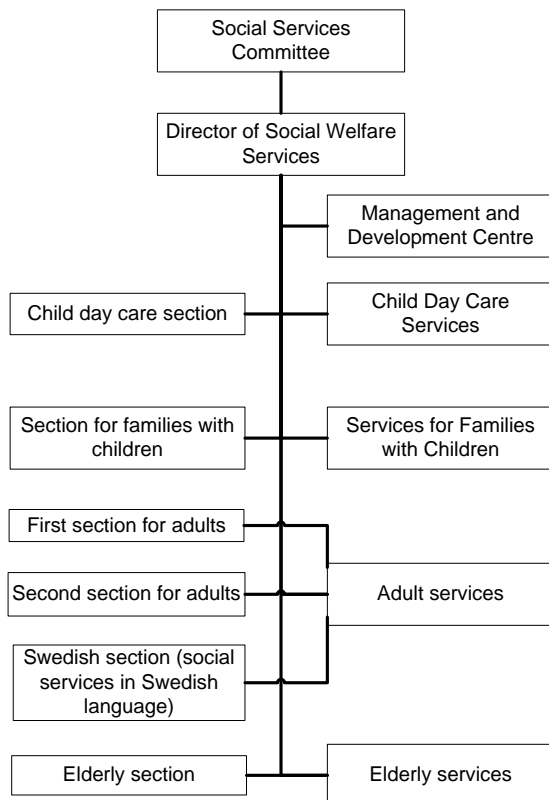


Figure 5 Organizational structure of the Social Services Department (Sosiaalivirasto, 2008)

The experiences reported in this study are related to the following services provided by the Social Services Department:

- child day care (cases in services provided in both official languages of Finland: Finnish and Swedish)
 - around 300 centers administered in 200 units and in 20 administrative regions
 - fewer than 10 centers provide services 24h/day
- child welfare
 - eight foster homes
 - some units provide extra services such as school and care for the disabled
- elderly services
 - 13 units of sheltered accommodation also providing care services
 - four service centers for clients needing special care
- disability services
 - seven day activity units for the mentally handicapped
 - services available only during normal office hours

These services, as a whole, represent roughly the half of the organization's own service production (excluding allowances) in Social Services Department in terms of costs. All the services are by nature classic services with close interaction between service the provider and the customer. They may also be characterized as high-touch services (Grönroos, 2000), which refers to the employee-intensity in the provision of services. The case services had different status in terms of measurement

in general. Some of the cases had a wide variety of measures including aspects of service quality whereas some services had, in practice, only the ‘common measures’ (used in practically every operation of the department) in use. The culture related to measurement was also more advanced in some services in comparison to others.

2.4 Research methodology

2.4.1 Research approach

The purpose of this section is to examine the overall research approach of this dissertation. More specific methodological issues are described in the next sub-section and in the individual articles. There are many factors affecting the choice of research approach. Most obvious one is the objective of the research in question. However, factors such as access to research data, resources available and the background of a researcher also influence the choice.

Studies on management typically try to understand and improve the performance of a business (Gummesson, 2000, p. 5) which is also the case in this study within the public sector context. Research on business economics can be characterized in many ways such as hermeneutical, positivistic, qualitative, quantitative, descriptive and normative. Different characterizations are often presented as contrasts to each other, although they are not necessarily mutually exclusive but complement each other. For example, a study is often not purely quantitative or qualitative.

A key difference between hermeneutic and positivist research is that positivistic research typically concentrates on description and explanation, whereas hermeneutic research deals with understanding and interpretation (Bryman and Bell, 2007; pp. 16–18; Gummesson, 2000, p. 178). Positivistic research has the ideal of objectivity whereas hermeneutic studies typically have more subjective features.

Qualitative research is typically related to hermeneutics, whereas quantitative research often represents positivistic thinking (Gummesson 2000, p. 178; Olkkonen, 1994, p. 35). Qualitative research is a kind of umbrella concept which may include many different kinds of studies (Tuomi and Sarajärvi, 2009, p. 9). Qualitative research emphasizes words instead of quantification in the collection and analysis of data and typically takes an interpretive approach (Bryman and Bell 2007, p. 402; Green et al., 2002). Qualitative research is very useful in the field of managerial research (Gummesson, 2000, p. 1). Qualitative research strives for a comprehensive picture of the research object with in-depth examination. Conversely, quantitative research includes examination of numerical data with an objectivist conception of social reality (Bryman and Bell 2007, p. 154; Green et al., 2002).

Descriptive research aims to illustrate certain phenomena, e.g. by creating concepts, describing processes, classifying phenomena and presenting correlations. Normative research, which has its roots in design sciences, aims to identify results that may be used as instructions for developing operations or designing new constructs. (Olkkonen, 1994, p. 44). In practice, constructs with a normative nature should be tested in order to validate them (Kasanen et al., 1993).

Case research is a fairly commonly used approach in the field of business economics. It has roots in social sciences and especially in ethnographic studies and anthropology (Voss et al., 2002). Case studies are often used to create theoretical constructs and propositions from case-based, empirical evidence (Eisenhardt, 1989). Case study is an all-encompassing approach which covers the logic of design, data collection techniques and approaches to data analysis (Yin, 2009, p. 18). Diverse (both quantitative and qualitative) methods and data may be used in case studies (Gummesson, 2000, p. 83; Meredith, 1998; Yin, 2009, p. 19). For example, they can include questionnaires, interviews, interactions, informal conversations, meetings, archival data, business plans, organizational charts, survey data and observations (Eisenhardt and Graebner, 2007; Meredith, 1998; Voss et al., 2002). One challenge in case research is that it is often a time consuming way of working (Meredith, 1998; Voss et al., 2002; Westbrook, 1995).

Case study starts with a research framework and questions (Voss et al., 2002) and is especially appropriate for examining how and why questions (Meredith, 1998; Yin, 2009, p. 10). It is not possible to provide detailed and mandatory procedures for an individual researcher in a specific case (Gummesson, 2000, p. 181). According to Yin (2009, pp. 19 - 20), there are at least four different applications of case research. First, it may explain causal links in complex real-life interventions. Second, it can describe an intervention and the real-life context in which it occurred. Third, case research may illustrate specific topics with a descriptive mode. Fourth, case studies can shed light on situations in which intervention has no clear set of outcomes.

Two types of case research may be distinguished. First, a case study may work towards general conclusions from a limited number of cases. Second, a case study may attempt to achieve specific conclusions related to a single case because it is of particular interest. (Gummesson, 2000, p. 84) Single cases offer an opportunity for in-depth investigation (Voss et al., 2002). Single cases may enable the creation of complicated theories since theory can be fitted exactly to the many details of a particular case (Eisenhardt and Graebner, 2007). They have, however, limited generalization of conclusions (Voss et al., 2002). Multiple cases can be used, e.g. for investigating the same issue in a variety of contexts. It is possible to achieve a more robust, generalizable and testable theory with multiple cases (Eisenhardt and Graebner, 2007; Voss et al, 2002). However, it is often impossible to present a complete and unbroken narrative with each case in a multiple-case study and therefore the depth of the study is impaired (Eisenhardt and Graebner, 2007; Voss et al., 2002).

This study can be characterized as *qualitative multiple case research*. There were many reasons for choosing this approach. The starting point for this study was the research and development project described earlier. In the initial phases of the project it became obvious that the existing literature did not, as such, support the development work carried out in the case organization. Hence, in addition to the practical solutions, there was a need to make a contribution to the research. The research project offered the much needed access to examine the research objective at hand (cf. Gummesson, 2000, p. 14).

As noted earlier in this study, examination of productivity and productivity measurement in public services is an extensive challenge. Qualitative research is suitable for examining complex issues and

phenomena (cf. Miles and Huberman, 1994). It was obvious that quantitative research would not, as such, achieve the objectives of this research and the requirements from the practical field (cf. Andriessen, 2004). Case research, as a flexible approach, has the potential for high validity with practitioners, who are the ultimate users of research (Voss et al., 2002). It was also the most obvious choice in investigating the research problem of this study. When discussing new managerial models and methods it is often hard to examine issues in a purely descriptive fashion. Therefore, this dissertation clearly also has normative features. Case research enabled the appropriate testing of the measurement methods in practice.

According to Eisenhardt and Graebner (2007), cases are not selected to represent some population. They are selected due to their suitability for understanding the issue at hand. Cases may be selected to represent unusual phenomenon, replicate other cases, contrast with other cases, eliminate alternative explanations and elaborate emergent theory. According to Voss et al. (2002), each case should be selected so that it either predicts similar results or produces contrary results, but for predictable reasons.

This dissertation examines one large organization consisting of several departments and services. In this large case organization five individual case services were examined. The same issues were investigated in various contexts of the same organization (replication). In this way, multiple cases provide the perceptions with more depth and credibility. The particular cases were chosen mainly because of the known productivity measurement challenges and also because of the significance in terms of resources used – productivity improvement could have great impacts. A practical reason for choosing these particular services instead of many others with measurement challenges was the representatives' willingness to participate. Even though all these cases were from the same department, some of them had quite different features (cf. child day care and elderly care). However, the main logic in case selection was replication in order to achieve better potential for generalization. It was estimated that experiences from these cases could be generalized not only to the same services in other organizations but also potentially to many other services with similar characteristics.

2.4.2 Data and methods

This section describes the methods used to gather and analyze the empirical data in this study. Figure 6 presents the main empirical research tasks carried out over three years. A literature review and conceptual analysis were the basis for the research work which continued throughout the empirical study. A detailed description of the methods used in this dissertation is presented next.

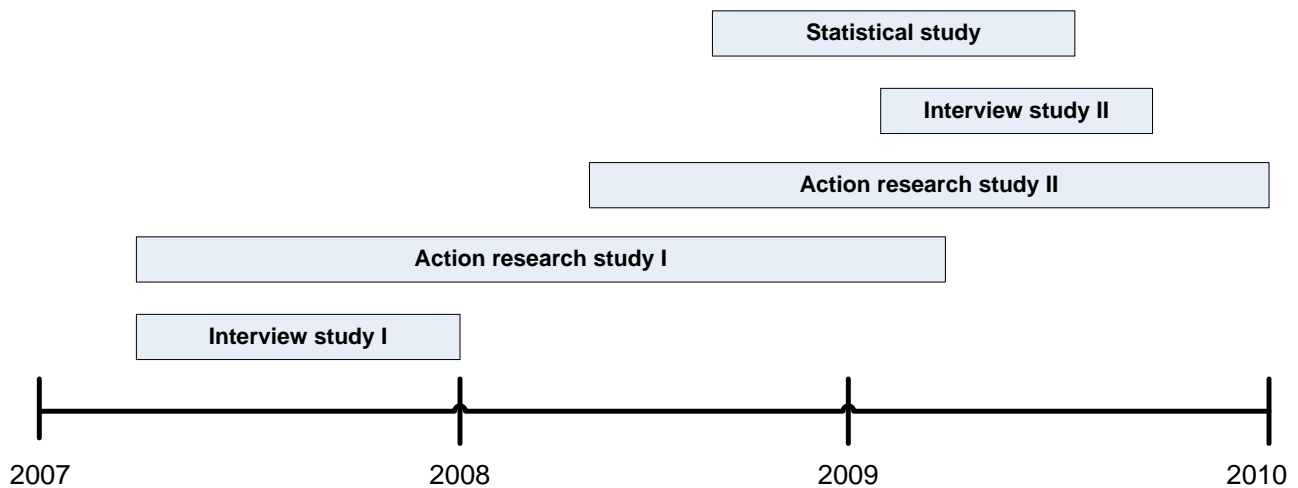


Figure 6 Use of empirical research methods

Interviews

Interviews are an efficient way to gather rich empirical data (Eisenhardt and Graebner, 2007) and often a part of qualitative studies. Interviews were used in different phases of the research work in this study. The first interview study examined productivity as a phenomenon, challenges and problems in productivity measurement as well as needs related to measurement. It was carried out at the beginning of the research work. The second interview study was conducted in order to evaluate the functionality of the measurement systems developed in the first action research study. In addition, it examined the factors affecting success in measurement system development. It was conducted in the middle of 2009.

The interviewees of the first interview study represented top-level decision-making in their respective organizations. In order to gain a comprehensive understanding of the issue, the interviewees were selected from all the key departments of the City of Helsinki. There were representatives of the technical services (e.g. Real Estate Department, Public Works Department, City Transport), sports and culture services, educational services as well as social and health care services. In total, there were representatives of 10 totally different public organizations (city departments). A total of 18 directors were interviewed individually. The interviewing was carried out by the author with the support of another researcher in five of the interviews. The duration of one interview was approximately one hour. In order to support the development work in the cases of Social Services Department, the four respondents representing the department were interviewed first. Transcriptions based on the perceptions and hand-written notes were used to document the data. Before each interview and during the analysis of the interview data, the researchers studied written documents related to the organizational structures of different departments and productivity measurement systems in use. The strengths of documentation as a source of evidence relate, among other aspects to stability (can be repeatedly reviewed) and exactness (contains exact names, references etc.) (Yin, 2009, p. 102). The purpose in this study was not only to improve the reliability of the analysis based on interviews but also to enhance the understanding of specific details related to measurement systems.

The first interview study contained 16 semi-structured questions which can be classified into three parts. The first part was related to the productivity phenomenon, e.g. what factors affect the productivity of the respondents' department and what is the significance of productivity from the point of view of the operations. The second part dealt with issues related to the current status of productivity measurement, e.g. the validity of the measurement system and the current utilization of measurement information. The third part concerned the needs related to productivity measurement: what purposes it should support, what are important criteria etc. These questions were chosen due to the goal of developing measurement. It was deemed important to understand not just the measurement object of productivity but also the needs related to measurement. On the other hand, it was also important to understand the current challenges and problems in productivity measurement in order to resolve at least some of them.

The 23 semi-structured questions of the second interview study can be divided into two parts. The first part dealt with technical factors of the measurement system developed including, e.g. issues related to the relevance, validity, reliability and practicality of the system. The second part was related to organizational factors in the performance measurement system development process and consisted of issues such as facilitation, the commitment of top and middle level management and also communication. The questions were chosen on the basis of the results of a literature review on the implementation of performance measurement systems.

Ten people working in the Social Services Department of the City of Helsinki were interviewed individually. The duration of one interview was approximately one hour. The respondents had been involved in the measurement system development project and therefore most of them knew the details well. The interviewees were managers on various levels, controllers and analysts. The interviewing was carried out by a researcher who had not been involved in the development of the measurement systems. The interview data was taped and transcribed.

Data reduction is the process of selecting, focusing, simplifying, abstracting and transforming the data from transcriptions. Coding is analysis involving differentiation and combining qualitative data. Codes are tags for attaching units of meaning to qualitative data. (Miles and Huberman, 1994, p. 10, 57) In the analysis of the data from interview studies I and II, the qualitative data was separated into individual classes (cf. Hirsjärvi and Hurme, 2004). These classes were mainly based on the results from the literature review. This may be referred to as a theory-based qualitative content analysis (Tuomi and Sarajärvi, 2009, p. 113). In the analysis of the needs for productivity measurement (part of the first interview study), the data itself was the starting point for the classifications. In the first interview study, the analysis was based mainly on subjective assessment by the author with support from another interviewer. In the second interview study, the analysis was based on the assessment of two researchers. In both studies, the results were sent to the representatives of the organization. They were discussed and commented in the meetings of the steering group of the research and development project. In this phase, no major changes to the analysis or results occurred.

Action research studies

The development of measurement in organizations always entails a change. According to Gummesson (2000, p. 15), the most advanced method for assessing organizational changes is action research. Action research is a specific type of case study in which the researcher is in active contact with the case organization (Coughlan and Coghlan, 2002; Gummesson, 2000, p. 116). It is a specifically appropriate method for examining research questions with a close relation to practice (Huxham and Vangen, 2003) which is also the case in this study. The work close to the managers provides a depth of understanding which is rarely possible with more traditional research methods (Westbrook, 1995). In action research the researcher is not just an external examiner in the case organization but also participates in the operations of the organization, for example as a part of a development project (Coughlan and Coghlan, 2002). Action researchers are external agents who facilitate changes in the organization studied (Coughlan and Coghlan, 2002; Gummesson, 2000, p. 120). On the other hand, they have to reflect and analyze the results in relation to existing research. Action research typically occurs in real time together with those who experience the issue directly (Coughlan and Coghlan, 2002).

A researcher may have different roles in the research work (Gummesson, 2000, p. 39). The role of change agent is the most appropriate description in this research as a whole, meaning that the research is action oriented and that researcher acts not only as an observer but also as an expert making recommendations. Observation is a key feature in action research. The strengths of observation relate to covering events in real time, covering the case context and gaining insight into interpersonal behavior and motives (Yin, 2009, p. 102). In participative observation social interaction is a key element in data gathering (Tuomi and Sarajärvi, 2009, p. 82). Consequently, action research provides such depth and credibility for investigating the objectives of this study that could not have been obtained merely by interviewing.

Specific types of action research include design-based research and constructive research common in the Finnish research tradition of business economics (Labro and Tuomela, 2003). Design-based research aims to develop knowledge that can be used in solving field problems (Andriessen, 2004; Van Aken, 2007). Designs may refer to entities such as actions, structures, processes or systems whereas design methods may be design processes, design roles and more specific methods for resolving certain design issues (Van Aken, 2007). A key phase in constructive research is the development of a new idea or managerial solution which should also be tested in practice (Labro and Tuomela, 2003). Constructive research is normative in nature (Kasanen et al., 1993).

A typical criticism of action research is related to the lack of rigor and labeling consultant projects as research. However, several differences between action research and consulting can be identified (Westbrook, 1995):

- consultants aim not to develop new theory but to report an application
- consultants share the single common goal with the organization whereas researchers pursue a wider goal, namely the discovery of new knowledge
- consultants specify the nature of the end result often before beginning the work while the researcher keeps an open mind
- consultants are likely to use established techniques rather than developing new ones

- consultants are often not interested in the journey to the destination, different challenges and success factors may go unreported

When comparing these characteristics with the objectives and results of this study, the nature of the research should be clearly identified. It should be noted that case studies with active researcher participation and intervention must be evaluated from the perspectives of both research and practice (Gummesson, 2000, p. 119; Lukka, 2000; Westbrook, 1995). More specifically, the effectiveness and validity of the solutions are not only evaluated by the researchers but also by the users in the field of application (Stam, 2007). According to Lukka (2000), constructive research may contribute to theory in two ways: 1) construction is as such so new that it may be used in identifying new means to achieve certain objectives, 2) constructive research may yield information for the purposes of development, illustration and theory testing.

One challenge of an action research project is an interactive and dynamic nature - it is impossible to plan or anticipate every detailed event beforehand (Coughlan and Coughlan, 2002; Gummesson, 2000, p. 119). The issue at hand often affects the way the research is carried out. In this study, the earlier knowledge about the phases and tasks in performance measurement system development (cf. Bourne 2003; Neely et al., 2000; Simons, 2000) served as a framework for carrying out action research.

In this study, the action research was carried out in interactive workshop events in which the new measurement systems were mainly designed. The length of one workshop was typically from two to three hours. In addition to the author (and sometimes a research colleague), 4-5 personnel (managers, analysts and accounting experts) working in the case organization attended the workshops. There were in total six working teams with their own workshop events. These interactive events facilitated a profound understanding of the procedures in the case organization and they were carefully documented. The author acted as a facilitator of the design process by presenting different measurement approaches from the literature, asking questions and directing the discussion as well as the design work. Written documentation was also utilized to learn more about current measurement systems (e.g. Balanced Scorecards in use), organizational structures etc.

Action research study I started in spring 2007 and lasted for 24 months. It involved two main tasks. First, operative productivity measurement systems were developed in the context of child day care (around 270 centers) and elderly services (13 units of sheltered accommodation). This work started with discussion on productivity phenomenon and the identification of the purpose and requirements for measurement. The results from the first interview study provided the perspective of top level management while the discussions in the first workshops yielded information on the same issues at the operative level. The second phase was to explore and evaluate various measurement methods in relation to the requirements identified. After that a specific method was applied. Finally, the measurement systems were implemented. Implementation involved test calculations, gaining the approval of the executive group, training sessions and development of an Excel tool supporting measurement. The development work required 10 workshop events in both case services.

The second task was related to the development of a method for the purposes of productivity measurement in the upper levels of the organization (see Article IV). The starting point for the work was the requirements identified for measurement. The actual design work was carried out by the researchers independent of the case organization. However, issues related to the testing of the method were discussed in three workshop events. The method was tested in spring 2009 with the measurement data (productivity change between 2007 and 2008) from all the child day care centers and units of sheltered accommodation for elderly people. Afterwards, the method and measurement results were evaluated by the steering group of the development project.

Action research study II started in spring 2008 and lasted for 19 months. Operative productivity measurement systems were developed in the context of child welfare (eight foster homes), elderly services (four service centers for customers needing much support) and disability services (seven day activity units for mentally disabled people). In addition, productivity measurement of Swedish language child day care (around 30 centers) was examined but no new measurement system was developed due to its similarity with Finnish child day care. The reason for dual examination and two working teams was mainly the separate administrations. In the second action research study, the same measurement method was applied as in the first one and therefore the development process was more straightforward. The development work required around seven workshop events in each case service. The main purpose of this study was to increase the understanding of applying the method by examining different services. In practice, this study yielded more information, especially on the measurement of quality-related aspects.

With both action research studies, the implementation of measurement systems were supported by training sessions addressed to managers, analysts and accounting experts working in the organization. The author participated in 12 training sessions by giving presentations and gathering comments on the measurement systems developed. Finally, the role of the meetings of the project steering group should be acknowledged. During the action research studies there were around 15 meetings in which the status of the development work was presented and evaluated. The steering group including three researchers (the present author and 1-2 professors), three representatives of the central administration of the City of Helsinki (chief planning officer, controller and budget director) and one representative from both the Personnel Department and the Social Services Department who mainly represented the view of top management.

From the point of view of practice, the results of the action research studies were new productivity measurement systems in all the case services as well as a method for aggregating the measurement results of the bottom-level units. From the point of view of the research the experiences and written documentation from around 50 workshops yielded information on the challenges in public service productivity measurement, needs for measurement, productivity phenomenon and the applicability of various measurement methods and frameworks.

Statistical study

Information provided by performance measures was used in the many test calculations carried out during the action research studies. Quantitative data gathered from performance measures was also examined in relation to the first main research question of this dissertation. Consequently, statistical analyses were carried out in order to learn more about productivity phenomenon at the operative level of public service organizations. The knowledge obtained also supported decision-making related to the actual measurement systems developed. The context for statistical analyses was child day care since it was already a familiar context and it had a specifically high number of similar service providing units. Therefore, it was a suitable context for statistically significant quantitative analysis. Pearson's correlation was chosen due to its practicality and suitability in cases in which the aim is to find a linear relationship between two continuous variables. Pearson's correlation is a method for examining the linear relationship between two interval/ratio variables. The key features of Pearson's r are:

- the coefficient lies between 0 and 1, which indicates the strength of a relationship
- the closer the coefficient is to 1 the stronger the relationship and vice versa
- the sign of the coefficient indicates the direction of a relationship (Bryman and Bell, 2007, p. 362).

Sample size and the size of the computed coefficient determine the significance of the correlation. The statistical significance of a correlation coefficient (based on a randomly selected sample) provides information about the probability that the coefficient is to be found in the population from which the sample was taken. (Bryman and Bell, 2007, p. 370) The p value gives the probability of error in the correlation coefficient.

To understand the productivity phenomenon examined more precisely, analysis of variance was applied. The method was used to determine the possible differences in performance factors between the means of three groups according to the size and productivity of the centers. To be able to apply analysis of variance, the dependent variable needs to be at least interval-scaled. Instead, the independent variable can be nominal-scaled. To find the groups between which the difference is significant Tukey's post-hoc test was used. F statistics is related to the explained variance (variation between groups) and error variance (variation within groups) and expresses the amount of explained variance in relation to the amount of error variance (Bryman and Bell, p. 371). The significance is explained by the probability of error as with correlation analysis.

2.5 Research structure

This dissertation comprises six research articles in which different perspectives and methods were used. Figure 7 describes the research question and also the organizational level (top vs. bottom; when specified) which the individual articles mainly address. The main contribution of this dissertation is based on the findings from the research papers, which are briefly summarized below the figure.

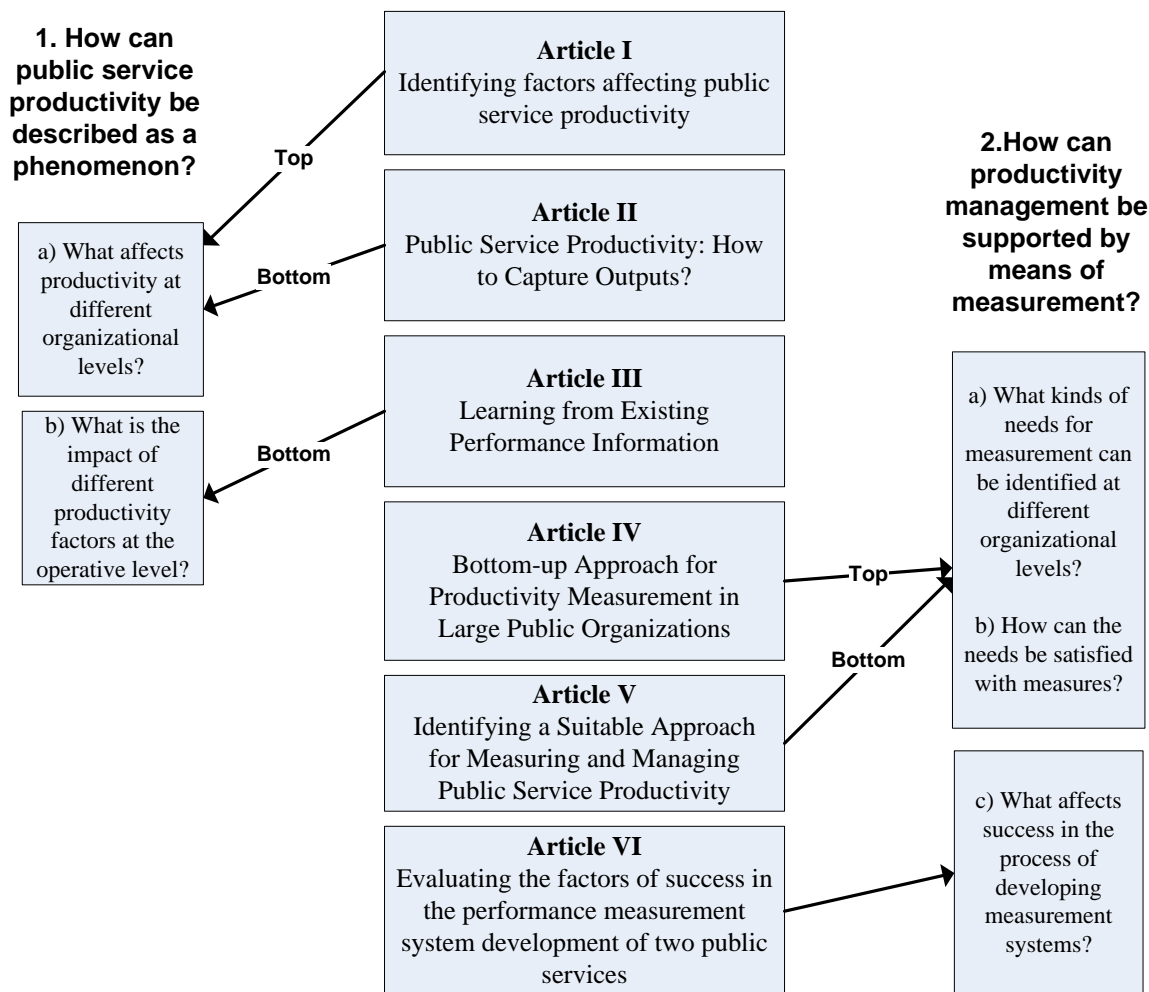


Figure 7 The composition of the articles

I Identifying factors affecting public service productivity

Jääskeläinen, A.

International Journal of Services Technology and Management, 2010, Vol. 14, No. 4, pp. 360–375.

In order to understand how to improve public service productivity, there is a need to analyze productivity phenomenon in the context. This information may also serve as a basis for developing appropriate measurement systems supporting productivity improvement. In comparison to performance-related studies, there are surprisingly few studies focusing specifically on productivity. The paper provides an understanding of the productivity phenomenon at the top level of a large municipal organization. The empirical analysis is based on interview study I. The factors identified as affecting productivity are analyzed by means of qualitative content analysis in relation to the results of a literature review. As a result, the paper presents a model of public service productivity presenting various classes and sub-classes for factors affecting productivity. The model can be used for identifying and analyzing factors affecting productivity in different contexts.

II Public Service Productivity: How to Capture Outputs?

Jääskeläinen, A., Lönnqvist, A.

International Journal of Public Sector Management (accepted for publication)

One of the challenges in the measurement of productivity in public services relates to the difficulties in capturing outputs properly due, e.g. to the intangible nature of outputs. A potential approach to overcoming some of the challenges is first to identify all the various output elements (both tangible and intangible). This article investigates various output elements as a part of managerial productivity measurement. From the point of view of this dissertation, the purpose of the article is to provide information on the productivity phenomenon at the bottom level of public service organizations. The empirical analysis is based on action research studies I and II investigating four different services. As a result, the article provides an extensive presentation of different output elements with examples from the case services. It is proposed as a first step in applying a disaggregated (component) approach in service productivity measurement.

III Learning from Existing Performance Information

Jääskeläinen, A., Kujansivu, P., Lönnqvist, A.

Proceedings of 5th Conference on Performance Measurement and Management Control, 2009, Nice, France.

Due to the proliferation of performance measures used, there is a need to improve their utilization. There is no sense in wasting resources on sustaining many measures if they are inefficiently used. There is not always a need to design new measures in order to satisfy needs for measurement. The paper has two objectives, first, to investigate how to better utilize the existing measurement information in monitoring key performance variables, and second (the main objective from the perspective of this dissertation), to better understand the dynamics of performance factors in public welfare services. The study is based on a case study approach utilizing various methods such as observations and workshops (first objective) as well as correlations and analysis of variance (statistical study related to the second objective). It focuses on the operative level with the empirical examination of one departmental service, child day care and its 270 day care centers. According to the results, disintegrated reporting is a key obstacle in utilizing existing measurement information. Statistical analyses are proposed as a potential tool for a better understanding of the characteristics of high performing organizations and also for identifying means to improve key performance objectives. The results from the statistical analyses indicate that public welfare services have many characteristics in common with classic services in general. For example, the role of customers is essential in productivity.

IV Bottom-up Approach for Productivity Measurement in Large Public Organizations

Jääskeläinen, A., Uusi-Rauva, E.

International Journal of Productivity and Performance Management (accepted for publication)

The productivity measures used in public organizations have been criticized for many reasons, such as the lack of relevance in managerial use. The present literature on public service productivity measurement is much emphasized in macro level measurement with poor linkage to operative management. The study proposes a new approach and method for measuring productivity in large public organizations. The main idea of the method is to aggregate productivity measurement information gathered at the bottom level of large organizations in order to provide productivity change indices for upper organizational levels. Hence, the measurement focus is more on top level examination. The research is carried out as a case study including interviews identifying needs for measurement (interview study I), and test calculations with the

measurement data from around 200 organizational units (as a part of action research study I). Based on the empirical examination the new method is argued to be more relevant than a more conventional counterpart for two key reasons: more detailed measurement information, flexibility to use different measurement methods in different services enabling the examination of administrative services with their own specific measures.

V Identifying a Suitable Approach for Measuring and Managing Public Service Productivity Jääskeläinen, A.

Electronic Journal of Knowledge Management, 2009, Vol. 7, No. 4, pp. 447–458.

The practical needs for productivity measurement at the operative level have not yet been systematically examined. This paper identifies and applies a productivity measurement method aiming to satisfy the information requirements of public managers, especially at the operative level (component measurement). The research is based on case studies in five services carried out as action research (action research studies I and II). In light of needs identified for measurement, the productivity matrix method is chosen from among three alternatives. The phases of designing a matrix-based measurement system are described according to the experiences in the case services. Finally, the advantages and disadvantages of the method are discussed.

VI Evaluating the factors of success in the performance measurement system development of two public services

Jääskeläinen, A., Sillanpää, V.

Proceedings of 10th EURAM Conference, 2010, Rome, Italy.

According to the literature, there are both technical (e.g. validity of measures) and organizational factors (e.g. committing employees) affecting success in measurement system development. However, the current knowledge of the impact of these factors in specific settings is limited. The specific setting of this study is related to a specific purpose (productivity management) and an empirical context (public welfare services). Specific factors affecting the success of measurement system development are examined based on the interview study II in two case services of the action research study I. From the perspective of the dissertation the contribution of this paper is twofold, first, it provides a user's view on the functionality of the matrix-based measurement systems developed, second, it describes the tasks and steps that supported success in the measurement system development of the case services. It is concluded that many of the known challenges in developing productivity measurement in large public organizations may be avoided by approaching the issue organizationally bottom-up. The representatives of the case services regarded the productivity matrix as a suitable tool for productivity management at the operative level.

Co-operation is a typical way of carrying out research. Four of the papers of this thesis were written in co-operation with other authors. The Table 5 describes the role of the author of this dissertation in each of those papers.

Table 5 Role of the author in articles with more than one author

Article	Role of the author
II Public Service Productivity: How to Capture Outputs?	<ul style="list-style-type: none"> - Designed the study together with the co-author - Facilitated the action research study - Analyzed the results together with the co-author - Wrote and reviewed the paper together with the co-author
III Learning from Existing Performance Information	<ul style="list-style-type: none"> - Designed the study together with the co-authors - Gathered measurement data together with the second author - Analyzed the results of the statistical study together with the co-authors - Wrote and reviewed the paper together with the co-authors
IV Bottom-up Approach for Productivity Measurement in Large Public Organizations	<ul style="list-style-type: none"> - Designed the study - Designed the measurement method together with the co-author - Carried out the testing and evaluation of the measurement method - Wrote the paper - Reviewed the paper together with the co-author
VI Evaluating the factors of success in the performance measurement system development of two public services	<ul style="list-style-type: none"> - Designed the study - Planned the interviews - Analyzed the interview results in co-operation with the co-author - Wrote and reviewed the paper together with the co-author

3 RESULTS

3.1 How can public service productivity be described as a phenomenon?

3.1.1 What affects productivity at different organizational levels?

At the conceptual level, there is no need for a different productivity definition in the context of public services. Public service productivity can be defined similarly as productivity in general – as a relationship between output and input. In operationalizing the productivity concept many questions arise: how to define outputs and inputs or what are the factors that should be stressed in identifying means to improve productivity? Analysis of productivity phenomenon is important in understanding why certain functions or units have good or poor productivity. It is also beneficial when developing measurement and management systems supporting productivity improvement. As a part of the interview study I, top-level directors were asked what in their view affects the productivity of public organizations. For the data analysis, factors were divided into individual classes based on the literature of service productivity and public service performance. As a result, a model of public service productivity was constructed (Figure 8).

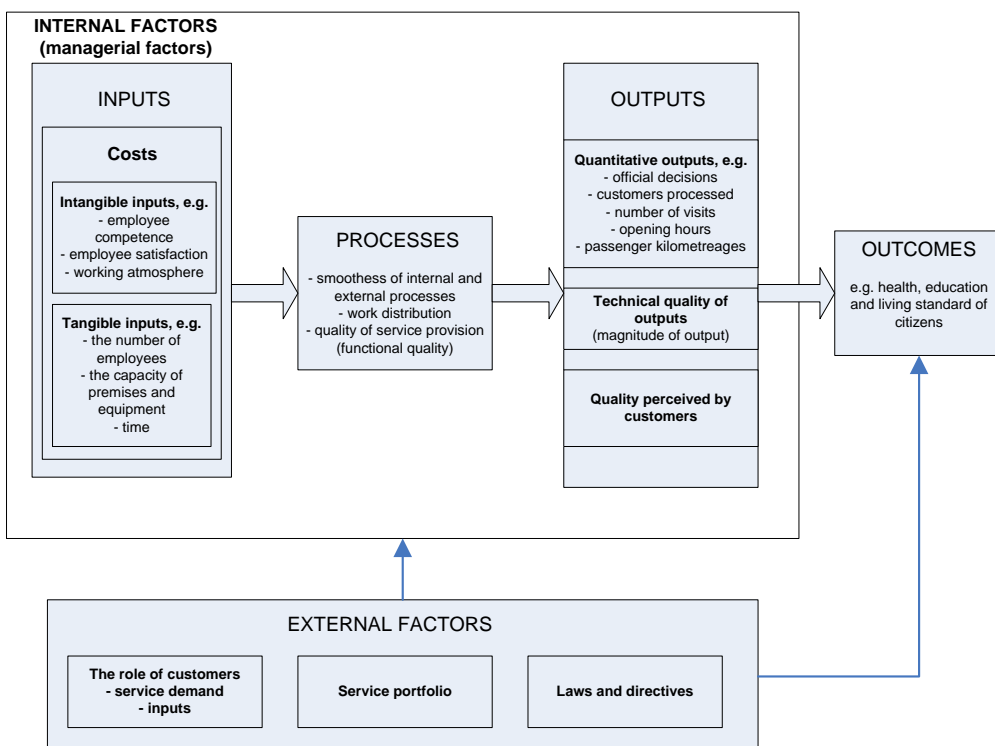


Figure 8 A model of public service productivity (Article I)

According to the interviews, the key classification of productivity factors was related to external and internal (managerial) factors. Externality refers to factors beyond the control of the managers of public organizations. Even though the distinction between external and internal productivity factors is not unambiguous, it may still be argued that some factors are easier to affect than others. This depends, e.g. on the decision-making power of a manager or director. The role of customers can

sometimes be affected by influencing the timing of demand or by increasing self-service. However, the demand may be heavily dependent on uncontrollable factors such as the economic situation. Other external factors include service portfolio: what individual services have to be available for the public, as well as laws and directives such as the definition of the ratio of customers to employees. These issues are mainly in the control of politicians. Political factors have sometimes been regarded as productivity barriers in the public sector since political decision-making is not necessarily based on productivity considerations (Ammons, 2004). However, political decisions can also be essential in ensuring that outcomes of services are desired. From the managerial perspective, external factors provide the frames for productivity improvement. They have to be acknowledged and anticipated despite limitations in control.

Internal productivity factors are controllable by managerial means and therefore highly relevant from the perspective of measurement and management systems. They consist of inputs, processes and outputs (cf. Stainer and Stainer, 1998). The interviewees identified both tangible and intangible inputs. The employee-intensity of public services was apparent in the responses. Intangible inputs were related to employee competence (sometimes referred to as input quality, Vuorinen et al., 1998), employee satisfaction as well as working atmosphere. There could also be other intangible inputs related, e.g. to relationships between different stakeholders or the values and culture of an organization (cf. Kujansivu, 2008, p. 7) but these were not mentioned by the interviewees. Tangible inputs were related especially to the number of employees, capacity of premises and equipment as well as time spent. From more conventional inputs, material and energy were not mentioned by the interviewees. As a whole, inputs were often related costs used to provide outputs (cf. Stainer and Stainer, 1998).

Existing inputs can be transformed into outputs both efficiently and inefficiently. This transformation may be related to various operations and processes in organizations. The smoothness of both internal (e.g. between administration and service providing units) and external processes (e.g. between departments) was highlighted by the interviewees. The role of rational work distribution was also regarded as an important factor affecting productivity. As discussed in the literature review, outputs are an important and challenging aspect in the examination of public service productivity. The model proposed distinguishes three output perspectives. First, output quantity relates to traditional quantitative outputs of public services. Aspects related to output quality were stressed by many of the interviewees. In the model, output quality aspects are distinguished by using Grönroos' (2001) service quality categorization into technical output quality (may be examined in a largely objective manner) and quality perceived by customer in relation to pre-conceptions (examination mainly subjective).

Even though outcomes are not usually included in productivity examination, it is clear that productivity should not conflict with effectiveness (cf. Linna et al., 2010). The role of outcomes should be acknowledged, since they represent the ultimate goals for providing public services. Outcomes of services are often also affected by factors external to the perspective of the organization's own decision-making (cf. Hill, 1977; Fletcher and Snee, 1985). Even though service outputs are provided as planned, some other factors such as customer's own decisions may affect the outcomes.

Challenges in service productivity measurement have sometimes been related to the intangible nature of service outputs (e.g. Flipo, 1988). In Article II, outputs were specifically examined from the perspectives of intangible and tangible output elements. Adapting the classification of that study, quality perceived by customers and outcomes include intangible elements whereas other output-related factors are tangible. According to the experiences of the study, the identification of such intangible elements is not the main problem – measurement causes much of the challenges.

Factors affecting productivity were also examined in the both action research studies (see Articles II and V). The key purpose was to identify relevant measurement objects for the measurement systems developed. A specific examination was related to various output components of the services. In order to facilitate the discussion in the workshops of different case services, drivers of productivity from the perspective of cost-efficiency and quality of services were identified. These two productivity-related goals were familiar to the case representatives and have also been mentioned in the literature (Faucett and Kleiner, 1994; Hodginson, 1999). When examining the results it should be noted that existing measures, as well as the purpose of developing measurement likely had some effect on the discussion. The responses are therefore possibly not as broadminded as in the interview study presented above. The results from the workshop discussions are summarized in Table 6, which is based on the classifications of the public service productivity model. No significant differences in different services were identified and the results are therefore examined jointly.

Table 6 Factors affecting productivity at the operative level

Productivity-related factor	Examples identified in the case services
Tangible inputs	<ul style="list-style-type: none"> - number of employees - premises - number of places for care - time
Intangible inputs	<ul style="list-style-type: none"> - working atmosphere - employee competence - qualification and education of employees
Processes	<ul style="list-style-type: none"> - smoothness of work processes - work distribution - prioritization of operations
Quantitative outputs	<ul style="list-style-type: none"> - number of care days - number of residential days - number of customer transportations
Technical quality of outputs	<ul style="list-style-type: none"> - differences in quantitative outputs caused by different age and special needs of customers - differences in quantitative outputs caused by different condition of customers - safety of customers - range of activities offered - availability and location
Quality perceived by customers	<ul style="list-style-type: none"> - customer (direct and indirect) satisfaction - preconception of the reliability and quality of service provider - personalized service
Outcomes	<ul style="list-style-type: none"> - children's development according to educational plans - successful life of children in the future - improvement in physical and mental condition of elderly people - ability of handicapped people to carry out work assignments
External factors	<ul style="list-style-type: none"> - customer structure - changes in customer demand - regulations on the number of employees per customers - level of rental costs - salary level of employees

Tangible inputs were similar to those identified by the representatives of top level decision-making. The role of technology or equipment was not specifically highlighted, possibly due to the employee-intensity of the services examined. The number of places for care was a more detailed definition of premises as an input. It was related to an existing measure in use. Time as an input was considered to be related especially to the specific needs of various customers. Intangible inputs discussed were also similar to those identified by top management. Along with working atmosphere, employee competence was highlighted in many of the case services. In addition, qualifications and formal education were regarded as important in ensuring the quality of services. These were also related to employee costs – qualified employees are paid better.

Many factors were related to actual operations and processes in which inputs are transformed into outputs. Employee sickness absences hinder the utilization of inputs and may also affect the quality of services provided. Employee turnover may have similar effects since it takes time and effort to achieve productivity potential in a new workplace (Kransdorff, 1995). In order to make the most of the potential of employee resources, various work assignments have to be allocated to the most suitable personnel. In addition, the work has to be planned in such a way that the most essential assignments are carried out first.

Like the directors interviewed, the representatives of the operational level identified quantitative outputs which were related to existing productivity measurement system. There was typically one main quantitative output representing the most of the operations in specific case units. Other quantitative outputs provided along with the main one were related, e.g. to the transportation services for handicapped people and day activities provided in sheltered accommodation for elderly people.

In all of the case services, differences in the technical quality of quantitative outputs were identified. In child day care, different ages and special needs of children affect the content of outputs. In elderly care and the care for handicapped people, the physical and mental condition of customers may vary, which is clearly a factor related to the technical quality of the quantitative outputs. Other factors related to outputs offered were also discussed that can be regarded as aspects of technical quality. Safety of customers was related e.g. to the existence of plans for fire and rescue measures as well as means (fences, locks etc.) to prevent customers from absconding from the premises. For example in the units of sheltered accommodation for the elderly as well as in day activity centers for handicapped people the variation of activities provided (e.g. fitness facilities) is a factor that can be evaluated fairly objectively and may therefore be regarded as technical output quality. Similarly, the availability of services (e.g. opening hours or queuing time) and location of service providing units (decentralized vs. centralized) are factors of technical quality.

A traditional examination of service quality relates to the quality perceived by the customer which is examined subjectively in relation to pre-expectations (Grönroos, 2001). In public services, the pre-expectations can, of course, vary, but may be related to the image of safe and reliable services without luxuries. In the case services, customer satisfaction was regarded as a direct result of outputs. A distinction should be made between direct (e.g. children in day care centers) and indirect (e.g. parents) customers since their satisfaction is possibly not congruent. Personalized care was regarded as a factor improving the quality of care, especially in elderly care and care for handicapped people. The existence of personalized care plans may be evaluated objectively but it does not ensure that the care is genuinely tailored. Therefore, personalized care can be regarded as quality perceived by the customer. The difference between subjective quality perceived by the customer and technical quality is not always clear. For example, customers may have their own views on the safety as well as the range of services offered.

Some of the quality-related aspects discussed in the workshops are more outcome-oriented. In the foster homes for children, quality was related especially to the success in achieving defined

objectives related to educational plans. In child day care, outcomes were related to the success in future lives of children which may be evaluated only after years or even decades. In elderly care, improvement in the physical and mental condition of customers was regarded to be an indicator of quality but at the same time a result from the care. An outcome example of the care for handicapped people can be related to the facilitation of customers' ability to carry out work assignments.

At the operative level, many factors may be regarded as external from the perspective of one's own decision-making power. According to the representatives of case units, there is only little possibility to affect the structure of arriving customers, since customers are assigned to the units by other personnel. The level of care needed as well as the number of customers may vary from time to time. There are also changes in demand that may be difficult to affect but can still be anticipated. For example, the demand for child day care is lower during the holiday season. Similarly, there may be systematically less demand on Mondays and Fridays. The demand may decrease more suddenly during flu epidemics. Furthermore, in some of the case services (child day care and substitute care for children), there were regulations related to the ratio between the number of customers and employees. In other services there were some instructions but no official regulations.

The final factors with a certain externality from the perspective of operative level managers were related to the use of monetary measures of inputs. Units are situated in different geographical areas, meaning that the level of rental costs per square meter may vary rather a lot. Similarly, in the day activity centers for handicapped people some units may be situated in such a way that transportations cost more in comparison to some other areas. Finally, the salary level of employees was considered partly difficult to affect. Experienced and qualified employees are paid more. This does not necessarily mean, however, that more quantitative outputs can be provided. On the other hand, it could be assumed that this has a certain effect on the quality of services.

The model of public service productivity developed on the basis of interview study I seems also to be suitable for examining productivity at the operative level. It is not too detailed and therefore has the flexibility needed. Individual factors affecting productivity may vary in different services, even though a certain analogy is apparent between the services examined in this study.

It could be summarized that both in the interviews and the workshop discussions the examination of productivity was at times very broad. The results described were based mainly on open and non-restricted views of the representatives of organization studied. Productivity was seen as an ultimate objective which can be achieved in many ways. A question that rises is: what the difference between such investigation and the general examination of public service performance (cf. Table 2)? At least three differences can be identified. First, there is more explicit linkage between productivity aspects affected by the view of the production process. Second, the role of customers and other external factors is specifically included. Third, the quality of services is included with various perspectives.

A comprehensive examination of productivity phenomenon can be useful in managing productivity and performance in general. However, in order to design specific systems for productivity measurement there is a need to prioritize and distinguish productivity from other performance

aspects. One of the most difficult issues is related to the role of service quality. There is a fairly unanimous view in the literature that quality should be a part of the productivity examination of public services (e.g. Hodgkinson, 1999; Rosen, 1993). When including quality in a productivity examination differentiation between the concepts of productivity and effectiveness is problematic. Technical quality is possibly closest to the traditional meaning of the productivity concept. It should be noted that there is no need for public organizations operating with limited resources to provide 'over-quality'. Hence, a standard and satisfactory quality level should be defined in every public service. In light of the productivity concept, this standard could somehow be related to the maximum level of quality that the public (in general) would still be ready to pay for.

3.1.2 What is the impact of different productivity factors at the operative level?

Productivity phenomenon was examined above based on opinions and views of the interview respondents and participants of workshop events. However, there may be some unknown characteristics due to the complexity of the phenomenon in question. The available measurement information related to some of the productivity factors enables more systematic analysis of productivity with the possibility to confirm or question earlier understanding on the issue. Article III (statistical study) presents a detailed examination at the operative level of welfare service provision with measurement information from child day care services. Due to the existing measurement information used, the productivity factors examined are not systematically and exactly the same as in Table 6, even though many can be easily linked to them. Productivity was considered to be composed of two components: cost efficiency (total costs / quantitative outputs with a weighting capturing technical quality) and customer satisfaction (average result from a survey).

In traditional classical services with close interaction between service provider and customer, productivity has been argued to decrease immediately if inputs are not adjusted to the change in demand (e.g. Grönroos and Ojasalo, 2004). The results from the statistical study demonstrated that capacity utilization is essential from the perspective of productivity. The proper utilization of labor input and premises was linked to high cost-efficiency. However, no linkage was found between these factors and customer satisfaction.

Another perception from the results relates to the role of customers in productivity improvement (cf. Ojasalo, 2003). First, customers with specific needs for care seem to require more employees, also increasing costs. This was demonstrated in the analysis of this study with a specific customer group of S2 children whose first language is not Finnish or Swedish (official languages in Finland). Second, sudden change in demand (caused by a large number of absences of children) appears to have a negative effect on the utilization of premises (which in turn impairs cost-efficiency). Third, specific customer groups may negatively affect the overall customer satisfaction. Fourth, specific customer groups can also have an effect on the employees. According to the results of this study, there are more sickness absences among employees in units with high numbers of S2 children.

In earlier studies, employee sickness absences have been linked to poor productivity (e.g. Miller et al., 2008). In this study, no connection between sickness absences and cost-efficiency was found. However, a significant correlation between high number of short sickness absences and low customer satisfaction was found. This may be caused by the use of substitutes instead of familiar

nursemaids. Earlier literature has linked highly educated personnel to low number of sickness absences (e.g. Lehtonen, 2007). The findings of this study are similar: there is a connection between high competence and the low number of sickness absences of employees.

It has often been stated that high competence of employees has a positive impact on productivity (Penrose, 1995; Xu et al., 2006). Hence, it could be assumed that investments in training have a positive effect on productivity. In this study, the effects of training were studied during a time period of one year. At least with such a short time scale no positive effects from training were identified. On the contrary, productivity seems to be lower in units investing in training. This finding is somewhat confusing and should be examined carefully in further research.

An issue that has been discussed rather a lot without unambiguous result relates to the size of service providing units in the public sector. In large child day care centers it may be easier to optimize the use of employee resources but there is a risk of negative effects such as a spread of diseases (Laine, 2005). Also according to the results of this study, large size of centers seems to be a double-edged sword. Cost-efficiency appears to be better in large centers but at the same time customer satisfaction is lower. However, the findings of another study with the same empirical data but different statistical analyses are encouraging regarding the large size of centers (Jääskeläinen et al., 2010). According to the results of that study, the productivity of individual units may be high with many different combinations of productivity factors. More specifically, both large and small child day care centers may have high productivity. However, it seems that large centers may achieve high productivity more easily – there do not have to be as many children per employees as in smaller centers. In addition, it seems that customer satisfaction is better in large and productive than in small and productive centers.

Since many of the results from the statistical study are in line with previous studies, it appears that the child day care services studied have many characteristics in common with so-called classic services. With these types of services it is essential to pay attention to the utilization of service providing capacity, management of human resources and anticipating the requirements of customers. It is also important to pay attention to the size of units providing services.

3.2 How can productivity management be supported by means of measurement?

3.2.1 What kinds of needs for measurement can be identified at different organizational levels?

Challenges in productivity measurement were already discussed in Section 1.4.1. Many of those challenges presented in the prior literature were also identified in interview study I. A part of the study examined challenges and problems in the previous measurement system in use. In general, much of the criticism was addressed to unclear, rough and unconvincing information on productivity. The following issues were criticized most:

- Outputs are poor and they do not capture service content provided for customers. Changes in output quality are not properly taken into account (cf. Ojasalo, 1999, p. 59; Sherwood, 1994).
- Measures are not linked to operational goals (cf. discussion on different views of organizational objectives in Rantanen et al., 2007). For example, the goal of improving public health can be achieved and at the same time quantitative output figures of hospitals go down.
- Measurement results are affected by factors that are not under own control (political decisions, demand for services, change in operating environment).
- Factors affecting the productivity of large organizational entities remain hidden. There should be an opportunity to gain more detailed information on productivity (why figures are good or poor). The logic of measurement must also be transparent. In addition, the analysis of measurement results should be improved.

The purpose of measurement is a key factor that should affect the design of measurement systems. It is difficult to develop a measurement system that is perfect for many different purposes (Wisniewski and Stewart, 2004). Productivity measurement may have many different purposes depending on the level of examination and the organization in question. As a part of interview study I it was studied how directors representing top-level decision-making saw the potential purposes of productivity measurement. In general, it was considered important that productivity measures support managerial purposes. Of the potential measurement purposes discussed in the literature, external reporting and contract management were not emphasized (cf. Greiling, 2005, Rosen, 1993). On the other hand, some purposes were articulated in more detail. The most mentioned purposes were:

- The use of productivity measurement in connection with rewarding systems was most often mentioned. In some of the departments this was the case already. It was deemed important that productivity improvement was supported by incentives. Lack of monetary resources was regarded as a practical problem in using rewards for motivating. A connection between rewarding and measurement in general has been suggested in many public sector studies (Chan, 2004; Greatbanks and Tapp, 2007; McAdam et al., 2005).
- Allocating resources and budgeting were regarded as a traditional use of productivity measurement and also as important. A simplified example of this purpose is the increase of resources in services in which there has been an increasing trend in outputs and vice versa.
- Identification of concrete targets for productivity improvement was considered essential. In practice, this would require that measures provided detailed information starting from the operative level. It was also proposed that different services and units should be examined with separate measures.
- Monitoring of productivity trends was mentioned in several responses. It was specifically related to productivity reporting to the political decision-makers and the public.

The literature presents comparisons and benchmarking as one potential purpose of measurement. Benchmarking is the systematic identification of the best practices of other organizations or units which has been presented as a potential strategy to improve performance of public services (Coe, 1999; Folz, 2004). The use of productivity measurement in comparisons was examined with a separate question. In general, productivity comparisons were regarded as a useful purpose of measurement. However, productivity comparisons were also considered challenging since

measurement results are often not comparable (e.g. due to the different characteristics of various departments and units). The following comparison purposes were the most mentioned by the respondents:

- Internal comparison of services provided by the City of Helsinki was the most mentioned. It was already done to some extent in some departments. Internal comparison can be carried out by comparing different city regions as well as individual units. The most meaningful context of comparison was related to similar units providing similar services (cf. Ammons et al., 2001; Rautiainen, 2004), which was also partly related to rewarding in those units.
- Comparison of own service production with private (purchased production or in general) organizations in the same sector was discussed in several interviews. It was regarded as challenging, since private organizations often give no information on productivity figures and they may also not be comparable. In any case, it was deemed important to understand whether certain services should be privatized and whether the private sector uses some practices or technologies that are more productive.
- Comparison of own services with other Finnish municipalities was regarded an interesting but challenging use of measurement. Helsinki was considered a unique municipality in Finland in which the operating environment is exceptional and operations are organized differently. However, it was proposed that certain individual services (e.g. catering) might be more comparable.

At the end of the interviews, more detailed criteria for productivity measurement were discussed. Many of the respondents stressed the simplicity, concreteness and comprehensibility of the measurement methods used. If a productivity measure is too complex, employees will not understand how productivity figures can be affected. These arguments are in line with Riggs (1981, p. 579) who emphasizes the practicality of productivity measurement instruments in order to use them in managing personnel. The respondents also considered it important to find ways to motivate employees to improve productivity. In order to improve motivation towards productivity, one of the respondents proposed that productivity measurement could be offered as a tool for daily management.

The purposes and requirements of productivity measurement were also investigated in action research study I. The representatives of the case services (operative level managers) emphasized especially the following two purposes for productivity measurement. First and not surprisingly, the ultimate need was the usefulness of measurement in the productivity management of individual units. This was often articulated as information on all the key factors affecting the productivity of individual services, meaning in practice an extension from mere examination of inputs and outputs. The second purpose was related to the productivity-related benchmarking of units providing similar services. Consequently, the measurement system should treat different units as equally as possible.

Other needs identified were related to more specific measurement criteria. First, workshop participants felt that the specific characteristics of different services should be taken into account better than in the previous productivity measurement system (cf. Rosen, 1993, p. 55). Second, quality of services was regarded as an important aspect of productivity that had been ignored in the previous productivity measurement system. Therefore, capturing quality-related aspects was

regarded as important in order to achieve a fresh approach to productivity measurement. Third, controllability of factors measured was regarded as important. Factors considered as external from the perspective of decision-making of individual units should have little effect on measurement results. This criterion can be considered important especially if the level of productivity is examined. In productivity change measurement many external factors have possibly more a limited effect, since they (e.g. rental costs in a certain city region in comparison to others) are similar during a longer time period. Fourth, the measurement system should be easy to understand with the transparent logic of producing measurement results. Finally, a practical criterion was related to the use of existing measurement information. Many measures were used in the case services but not efficiently. Better utilization of existing information was also meaningful from the perspective of time and resources used. The development of many new measures with information systems, collecting of new measurement information etc. would require more time and resources than are available for development work.

It seems that there are many similar needs for measurement both at the top and bottom levels. Similar purposes for measurement were comparisons of units and identification of targets for productivity improvement. An example of a similar criterion for measurement is the practicality and simplicity of the measurement method chosen. There are also differences in measurement needs. Linking of productivity measurement and rewarding, or productivity trend monitoring were not highlighted at the bottom level. On the other hand, top-level directors did not find the chance to influence the measurement results problematic as did the representatives of operative level.

3.2.2 How can the needs be satisfied with measures?

In light of the needs for measurement identified by interviewing top-level directors, the following guidelines for productivity measurement development were defined and approved by the steering group of the development project:

- Productivity measurement should take into account the differences in services provided by the whole city organization. Different characteristics of services may have a role in how the concept of productivity is operationalized for measurement purposes. In order to capture the different features of services, there must be various ways for measuring productivity. Many of the challenges in measurement identified by the interviewees (e.g. output definition and the linkage between operational objectives and productivity measures) are caused at least partly by the fact that measures are not genuinely developed on the basis of various activities.
- Productivity measurement should provide more detailed information by examining smaller organizational entities with separate measures. This enables more meaningful productivity comparisons and also provides more information on the reasons for productivity changes. It may also be argued that productivity measurement becomes more relevant at the operative level and facilitates the development of measures capturing the specific features of various services.
- Measurement of outputs should be highlighted. Outputs should take into account the specific features and operative objectives of services, which means that the representatives of various services should have an essential role in measurement development.
- Productivity measurement should be more goal-oriented. Monitoring of productivity trends is not enough, there must be objectives related to productivity. This also means that measurement should be transparent and understandable enough.

It was decided that the various needs for measurement are best met with a measurement approach aggregating productivity information provided by component productivity measures used at the operative level. In this way it would be possibly easier to better satisfy the needs at the operative level. In addition, support for this approach was found in the existing literature (Boyle, 2006; Kald and Nilsson, 2000; Mammone, 1980; McLaughlin and Coffey, 1990; Riggs, 1981, p. 579). Several productivity measurement methods have been presented in the prior literature but no obvious solutions for operationalizing the approach were found. Existing methods typically first aggregate outputs and inputs and then examine the ratio between them at the level of organizational department or the whole municipal organization (cf. Coelli et al., 2005; Hackman, 2008; Simpson, 2009). Productivity information is therefore not available at the operative level. Consequently, a new bottom-up productivity measurement method was constructed (Article IV). Component measures designed for operative productivity measurement are the starting point for applying the method (see Figure 9).

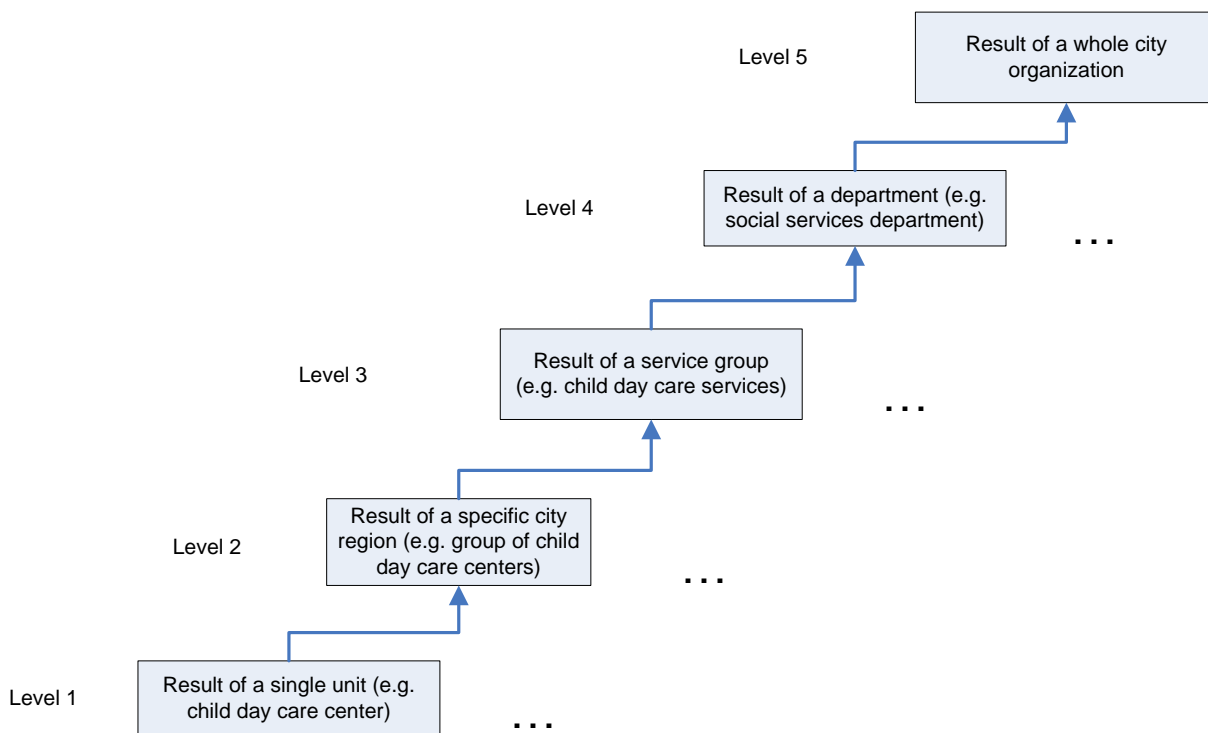


Figure 9 Principle of bottom-up productivity measurement

The following formula describes the calculations in aggregating the results from component measures of productivity. Productivity changes are examined since they are often considered essentially comparable even between different services (see e.g. Kangasharju, 2008, p. 186). Productivity change is measured with the ratio between productivity level in a particular year examined and productivity level in the base year of the whole examination period (e.g. 5-6 years). Productivity changes from individual units are aggregated by using the total costs of each unit per the total costs of the whole organizational level as a weight. The most accurate result is achieved if the weights (W_i) are updated every year, even though the cost structure of the base year may also be used.

Productivity change between the year examined and the base year = $\sum_{i=1}^{i=n} W_i * \frac{B_i}{A_i}$

Weight of an individual organizational unit = $W_i = \frac{C_i}{C_T}$

B = productivity value in the year examined

A = productivity value in the base year

n = number of organizational units

C_i = total costs of unit i

C_T = total costs of the whole organizational level

Sub-units at the operative level utilize an appropriate method for achieving the value for productivity change. The productivity change of individual organizational level is calculated by summing up the productivity changes of all the sub-units of the level of examination (see Figure 10). This value is used in order to obtain the value for the productivity change of the upper organizational level. When carrying out similar calculations in every organizational level it is possible to achieve one figure for the productivity change of the whole organization. Productivity of administrative services at each organizational level can be taken into account with separate measures. In the actual aggregations they are treated as individual sub-units.

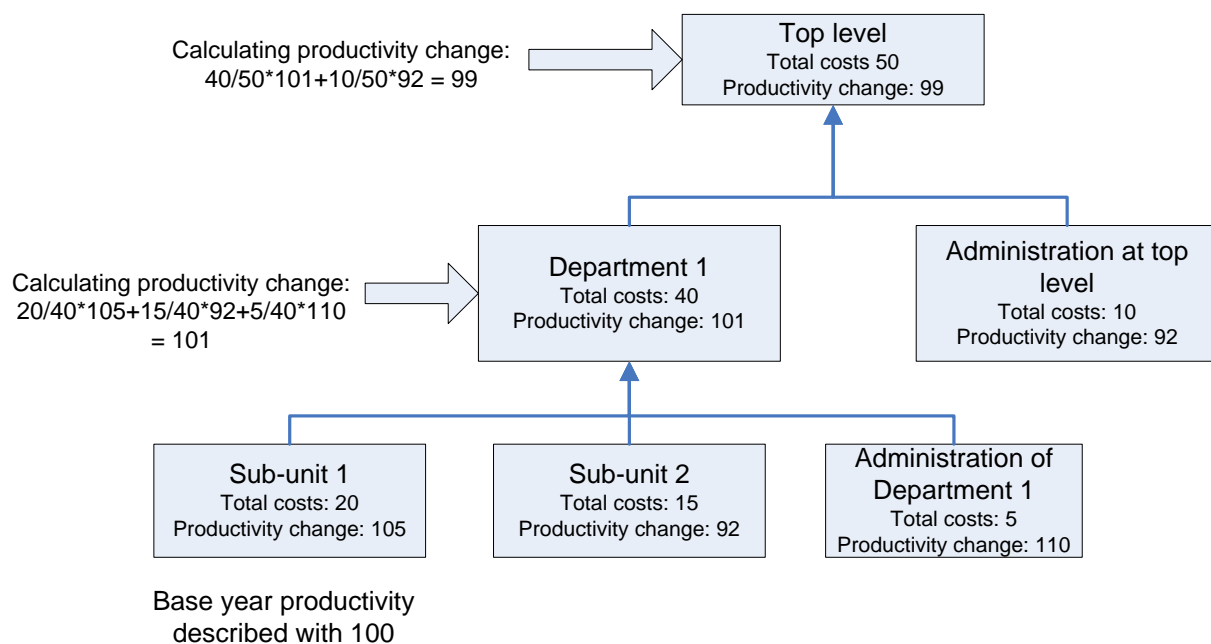


Figure 10 Example of calculations with the measurement method

The method enables the fulfillment of many of the measurement needs identified (e.g. gathering of detailed information and taking the characteristics of different services into account). It is possibly not better in allocating resources since separate information on quantitative outputs and inputs is not available at the upper levels. However, this information may be available from other existing

measures. At best, the method combines micro and macro level examination of productivity measurement. At the macro level the more detailed basic information provides more understanding of the reasons for productivity changes. At the micro level productivity information provides support for developing operations. However, the actual content of measurement as well as the quality of results achieved ultimately depends on the component measures used at the operative level. Hence, the main challenge is to develop appropriate component measures.

After evaluating three general alternatives for productivity measurement (see Article V), development of component measurement was carried out using two main paths in the case services of this study. First, productivity matrices for different services were designed (see examples in Appendix 1). In the initial stages of the design work (action research study I), the representatives of two case services regarded the matrix method as the most suitable for operative level purposes. It was considered that matrix enables a more comprehensive examination of productivity with measures related to many of the operative productivity factors. Other important aspects were that matrix provides information on productivity level (not just change) supporting benchmarking of units, illustrates results that are considered to be good, average and poor and also facilitates target setting. In addition, matrix gives one index from many different measures, which was considered important (ease of controlling measurement information), especially when many similar units provide the same services. Since the measurement scores regarding productivity level of similar units using similar matrices are essentially comparable, it was also perceived as a potential tool to be used for purposes of rewarding².

In interview study II (reported in Article VI) carried out after the design and implementation of matrix-based measurement systems in two case services of action research study I, the respondents evaluated the relevance, validity, reliability and practicality of the new measurement systems. They were considered relevant managerial tools at the operative level. Most of the respondents reported that systems covered important productivity factors identified (indicating validity) well. A validity-related issue that was not discussed in the interviews is how the weights and scales used in matrices affect validity. Weights were decided according to the opinions of the representatives of the case services (importance of factors from the perspective of productivity) and were partly strengthened with the results from the statistical study. The scales, in turn, were determined on the basis of extensive test calculations, which should at least partly eliminate validity-related problems. However, the role of these technical factors will eventually be revealed after the system has been in use for a couple of years.

The respondents considered the reliability of the systems fairly good, with only limited possibilities for the intentional manipulation of measurement results, despite the fact that some of the measures were subjective in nature. In any case, it was estimated that intentional manipulation of the results would eventually be exposed. Reliability challenges were linked to the data in information systems. Problems had been identified in the registration practices of measurement information (a task

² After the development project new matrix-based measurement systems were, in fact, about to be linked to the rewarding system, meaning that the productivity measurement purpose mentioned by most top directors was about to be fulfilled.

typically carried out by individual units). For example, there had been different ways of calculating quantitative outputs. With regard to the practicality of the systems, they were considered laborious at the time of the interviews since information systems provided only limited support. In addition, organizing the utilization of the systems (coordination etc.) was considered time-consuming in the initial stages of system use. However, the systems were considered to be logical and easy to understand with clear benefits for managerial purposes.

Measurement systems may also be evaluated according to the information they provide in relation to the actual perceptions of reality. During the workshops in both action research studies, many test calculations with the matrix were analyzed. Many differences were identified in the productivity factors of units providing similar services. There were also notable differences in the total productivity scores produced by the matrix tool. The representatives of the case services were able to link many such results to the characteristics of different units (e.g. location in the city) as well as various changes such as renovation and relocation of premises. Such results in line with the observations from the real life may be regarded as indicators of the appropriateness of the measurement systems developed.

The second development path of component measurement (carried out along with the development of the productivity matrices) was to design output weights for different services in order to better capture the technical quality of outputs. These weights were used in fixed unit cost measures (e.g. costs of a day in care) by multiplying quantitative outputs by defined weights. In order to keep the measurement systems practical, only one quantitative output per case service was used. Hence, in every service, the major quantitative output representing the most of the operations and inputs used was first identified.

New quality weights were eventually developed or existing ones applied in all the case services. In child day care there was already a weighting system in use based on the regulation. In service centers for the elderly RUG classification was applied. The classification is based on the evaluation by health care professionals working in the centers. In sheltered accommodation for the elderly the weighting system described in Table 7 was designed since RUG was not available during the development work. It is related to the service charges defined on the basis of the use of care in the units. A customer classified as receiving over 91 hours of care a month is defined to be equivalent to 6.3 customers taking only residential services. Similarly, in the care of handicapped people a classification system was developed based on the time used to serve customers with different physical and mental conditions. Condition is defined by the employees working in the unit.

Table 7 Output weighting system of sheltered accommodation for the elderly

Customer type	Output weight
Class 1 Resident using no care services	1.0
Class 2 Use of care services 1-10 hours/month	1.3
Class 3 Use of care services 11-30 hours/month	2.1
Class 4 Use of care services 31-60 hours/month	3.4
Class 5 Use of care services 61-90 hours/month	5.0
Class 6 Use of care services over 91 hours/month	6.3

These output weights capturing the technical quality of services were well received by the representatives of the case services, who commented that differences in customer structures were clearly visible in the measurement results. Therefore, the use of weights was believed to improve the comparability of the units. In addition, productivity could now be improved not only by decreasing input use or increasing the quantity of outputs but also by producing more demanding outputs (cf. Ilmakunnas, 2008, p. 3). Challenges in the use of such weights were related to the definition of functional weights on the basis of information that was as objective as possible. For example, in the case of care for handicapped people it was discussed how to avoid reliability problems related to subjective evaluation regarding the condition of customers carried out in different units. The practice of moving employees to work in another unit was tested with promising results. In light of the working experience of one week in another unit the external worker may better evaluate the condition of customers in relation to his/her own unit.

If weighted unit costs with fixed monetary information are used as component measures, productivity change can be calculated as follows (unit cost represents the inverse of output/input figure, base year described with 100):

$$100 * \frac{\text{Weighted unit cost of the base year}}{\text{Weighted unit cost of current year}}$$

Weighted unit cost was an essential part of all the matrix-based measurement systems. During the period of the empirical examination of this study, no decision was made as to which kind of component measure (weighted unit cost or matrix score) would be used in the aggregated productivity figures. According to the test calculations carried out with both alternatives, it seems that productivity changes produced by matrix measurement are more substantial. This is an obvious observation since the principle of matrix is different and it has a much broader perspective on productivity. It should also be noted that unit cost is continuous variable while matrix score is discrete due to the scales used. Hence the productivity changes calculated with the productivity matrix are not fully comparable with more conventional productivity measures. The criterion of comparability of productivity changes produced by matrices and more traditional productivity

measures is decisive in the eventual choice of component measures. In both alternatives, matrix-based measurement systems can be powerful tools at the operative level of productivity management.

Developing new measures and measurement systems is not the only option in improving measurement related to productivity management. Since many organizations have been applying performance and productivity measurement for years, there are often numerous measures already in use (Jarrar and Schiuma, 2007). Sophisticated measurement methods do not ensure their efficient use (Vakkuri, 2003). A general aspect regarding the development of measurement in the case services of this study relates to the use of existing measurement information – one reason for choosing a matrix approach to measurement was to better utilize existing measurement information.

In the development of matrix-based productivity measurement systems, much of the challenges and development effort was related to the measurement of service quality which has, in fact, been perceived to be one of the biggest challenges in performance measurement of public services (Poister and Streib, 1999). Even though considered important, there were sometimes no existing measures, nor even any clear definition for the quality of services. Customer satisfaction surveys addressed to indirect customers (e.g. parents of children) were sometimes used and included in the matrix despite the obvious limitations of such measures. In addition to the output weights described above, quality measures related to customer satisfaction (e.g. survey on customer satisfaction of handicapped people) and outcomes of care (e.g. survey on children's development according to educational plans) were specifically developed in services where no existing measures were available. To summarize, quality-related measures were the most tailored in every case service. However, many other existing measures were used as such or with small adjustments. Examples of such measures were utilization rate of premises, employee-customer ratio as well as percentage of sickness absences. Sometimes measures (e.g. person-years/places for care) used in other services were copied to another context. Hence, it may be argued that one of the benefits of a matrix tool is to give existing measures more significance by pre-analyzing measurement data. A substantial amount of data is compressed into one index, which is beneficial, especially when there are many units to be examined.

A separate study on using existing measurement information was carried out in the context of child day care (Article III). It was noted that a challenge in the use of measurement information was related specifically to disconnected reporting. For example, gathering all the measurement data (11 measures from various information systems) for a certain organizational unit required a substantial amount of manual work. Another perception was that there were measures not properly connected to managerial systems – they remained distant statistics for managers at the operational level.

The study demonstrated that statistical analyses may well be practical managerial tools for analyzing the service productivity of large public organizations if supported by appropriate information systems. For example, they may be used in order to understand what the characteristics of units with high productivity are or how different factors affect productivity (directly or indirectly). Two types of knowledge were provided by the analysis: totally new knowledge and knowledge supporting earlier assumptions. In Article III, rather simple statistical methods

(correlations and analysis of variance) were used. The potential of statistical analysis may be improved by using more sophisticated methods (such as SFA) of productivity analysis but the aspect of practicality should not be forgotten. In addition, the use of statistical analysis requires that an organization already has a comprehensive battery of measures in use. There must also be a large number of organizational units in order to have a sample size large enough for the analysis.

3.2.3 What affects success in the process of developing measurement systems?

In general, there are numerous studies on the process of developing performance measurement systems presenting guidelines and potential pitfalls (Bourne et al., 2000; Bourne et al. 2002; Lönnqvist, 2004; Rantanen et al., 2007). It has been claimed that there are both technical factors (validity of measurement system, functionality of IT systems etc.) and organizational factors (support from facilitators, communication with employees etc.) related to change management which can facilitate success. In the process of developing new measurement systems in the case services of this study, the intention was to utilize this existing knowledge in order to avoid at least some of the potential pitfalls. Factors affecting the success of measurement system development in the case services of first action research study were examined in Article VI (interview study II). From the perspective of creating new knowledge, the setting of this study was interesting due to specific features such as productivity as a measurement object and the context of the public sector. Among the managers of many public services productivity is not necessarily the first thing in mind and may also have a very negative image (cf. Laine, 2005). It was therefore valuable to assess the productivity measurement system development process, which can be regarded as fairly successful.

According to the results of interview study II it can be argued that the bottom-up approach chosen for the development of productivity measurement was a key facilitator of success enabling commitment of the operative level and satisfying managerial requirements. This concurs with earlier studies. For example, Kaplan and Norton (2001) stated that one of the main reasons for the failure of performance measurement system development projects is the sole emphasis on the top level of an organization. In addition, performance measurement systems are often designed by individuals other than their actual users (Vakkuri and Meklin, 2006). Cavalluzzo and Ittner (2004) have noted that performance measurement mandated organizational top-down by political decisions is likely to be symbolic with limited use and influence on internal operations. Lawton et al. (2000) argue that a top-down approach easily leads to biased performance examination. According to Lupton (1991), a change can be more successfully introduced from bottom-up than top-down. In relation to this, Rosen (1993, p. 41) and Chan (2004) have claimed that it is vital to crush the resistance to change among middle management in order to achieve success in productivity and performance development efforts. This was also a notable factor in the case services of this study. According to the results of Article VI, participation by the representatives of operative services in the development of measurement systems was important in overcoming resistance to change. This concurs with perceptions from earlier studies (Likierman, 1993; Ukko, 2009, p. 51). For some reason this seems not to have been the case in the public sector productivity examination.

More specifically, Lancer Julnes and Holzer (2001) have noted that when the policy of using performance measurement is based on internal requirements (e.g. managerial needs), measures are more likely to be adopted. In this study, capturing service quality was an essential aspect regarded

as an important criterion for measurement (and eventually captured with the measurement systems). Hodgkinson (1999) has also emphasized this by arguing that ignoring quality in the evaluation of public services may impair the motivation of employees to improve productivity.

Many of the factors that facilitated the success in the measurement system development of this study, such as identification of managerial requirements, were accomplished early in the design process. Both technical and organizational factors were essential but it may be argued that organizational factors, such as communication with top and middle management, were even more important in the specific setting of developing productivity measurement for public services.

In addition to the views of the representatives of the case services obtained through interview study II, experiences based on the perceptions of the author during both action research projects provide an understanding of the possible causes underlying the success of the development of productivity measurement systems in the public sector. These perceptions are more related to the technical factors of measurement system development.

First, the identification of managerial needs and the analysis of the measurement object (e.g. service quality and its role in productivity) were essential starting points for the work. When developing measurement systems, it is all too easy to short-cut and start the discussion from technical measures without asking why to measure. Bourne et al. (2002) have also noted that one of the characteristics of failed measurement system development projects is that the reason for measurement development is related to the need to improve measurement instead of the need to improve management. Detailed examination of the factors affecting the main measurement object (productivity) was also beneficial in this study. This was facilitated by the use of maps similar to strategy maps (cf. Kaplan and Norton, 2004) describing cause and effect relationships between factors affecting productivity. Norreklit (2000) has also stated that a proper understanding of the cause and effect chain is essential for functional performance measurement.

Second, the prioritization of measurement objects and the ambition to keep things simple facilitated the design work and possibly also improved the final result. Many studies have identified an essential problem of measurement systems in the public sector: they try to measure everything for everybody. This may mean that measurement systems include a substantial number of measures, which, in turn, impairs the usefulness of the systems (Atkinson et al., 1997; Bierbusse and Siesfeld, 1997; Kaplan, 2001; Rosen, 1993, p. 73). An underlying problem is the difficulty of identifying one ultimate objective (like profitability in the private sector) of public organizations (Rantanen et al., 2007). In the cases of this study, the ultimate objective was more focused (productivity) than overall organizational performance. This meant that there was a clear reason for prioritization, which served to facilitate the task. Still, there was an almost endless discussion on the factors regarded as essential and likely to affect productivity figures. For example, employee turnover and the number of short care periods were considered to affect the productivity of elderly care. As a result of prioritization and challenges in measurement many of those factors were excluded from the measurement systems. Background information related to measures is almost always necessary and it may also be provided without any specific measures. This kind of information should be systematically documented and generally accepted.

There was also an ambition to keep things simple enough in order to achieve results with reasonable use of time and resources: not everything was changed at once but existing information and systems was utilized whenever possible. This meant that some potential factors to be measured were initially omitted from the measurement systems due to difficulties in measurement (e.g. satisfaction of children in day care centers) or the potentially substantial work needed to adjust existing information systems in order to provide the information needed (e.g. internal employee turnover). This decision is also supported by the perceptions by Chan (2004) and Rosen (1993, p. 72). In addition, McCunn (1998) has noted that striving for perfection hampers the achievement of success in a measurement system development.

Finally, it should be noted (as also indicated by the responses in interview study II) that when this dissertation was being written the work related to measurement systems was not yet totally over – actual benefits (and eventual success) may only be realized if the final steps are carried out properly. In the actual use of matrix measurement many details, such as target values, frequencies in measurement and responsibilities (see e.g. Neely et al., 1996, p. 64) have to be clear. Many of these issues were already determined during the action research studies with convergent principles (e.g. matrix score is officially calculated once a year). A more difficult question (closely related to rewarding) that was yet to be decided was the setting of targets for the total score of the matrix and for the score of individual measures. Should there be individual targets for each unit and how should the targets be determined? The matrix method has ‘built-in’ targets (score 10 for each measure), which however, are not realistic short-term targets for many of the units. There is a risk that poorly set targets for individual units (e.g. targets that are too easy to reach) may even impair the overall performance of public organizations (cf. Rajala and Meklin, 2008, p. 108). The discussion on this issue is beyond the scope of this research due to practical limitations. Further research should be carried out after there are more experiences of the use of measurement systems.

3.3 Summing up the results

This study examined productivity measurement and management in large public service organizations with two viewpoints related organizationally to top (e.g. municipal department or a whole municipal organization) and bottom levels (e.g. administrative unit providing a specific service). Figure 11 summarizes some of the key results of this study.

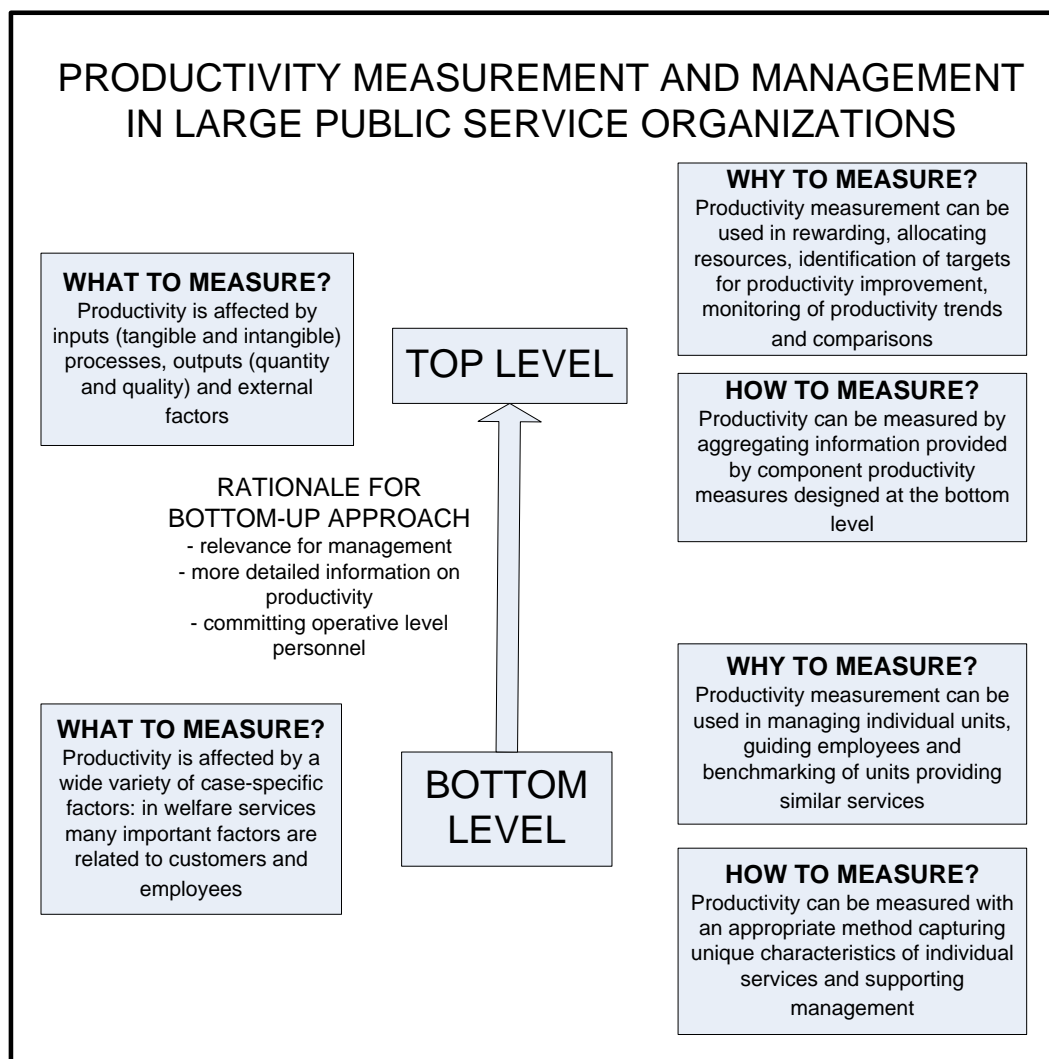


Figure 11 Summary of the results

At the top level, productivity phenomenon can be examined from various perspectives. Not all the productivity factors can be affected even at the top managerial level since there are externalities such as legislation. It is necessary to examine not only tangible inputs but also employee satisfaction and other intangible input factors possibly affecting productivity. Similarly, outputs comprise not only quantitative factors but also qualitative factors (technical quality and quality perceived by customers). In order to better understand productivity phenomenon, it must be examined at the operative level in which there may be a variety of more detailed case-specific factors. The classification used at the top level, however, can also be utilized in analyzing productivity factors at the bottom level. In the welfare services examined in this study, many important factors affecting productivity are related to customers and employees. Customer demand (e.g. number of customers) may vary, meaning that service providing capacity needs to be adjusted in fluctuations in demand. This requires careful anticipation with the information available. In addition, customer needs may be different due to a certain level of customization. This causes problems in productivity measurement, since it may be difficult to define standard outputs. This is also an issue that should be taken into account in the division of labor in order to have resources available in those units and subunits where there are the most demanding customers. Employees are a key input resource. The close linkage between employees and customers in many welfare services

should be acknowledged when productivity is discussed. For example, competencies and satisfaction of employees may affect the efficiency of work as well as have a crucial impact on the service provision and quality perceived by a customer. This research showed that employee sickness absences may impair customer satisfaction.

Many different purposes for productivity measurement were identified by interviewing directors representing top-level decision-making. In addition to more conventional measurement purposes related to allocating resources, budgeting and monitoring productivity trends, it was also noted that there is a need to link productivity measurement to rewarding as well as to the identification of more detailed targets for productivity improvement and productivity comparisons. At the bottom level, the investigation of needs for productivity measurement based on the workshop discussions was more focused on detailed criteria for measurement such as capturing of quality aspects and the controllability of the factors to be measured. In addition to the support for productivity management at the operative level, productivity measurement was related to benchmarking of similar units.

In this study, productivity measurement was investigated with an approach measuring productivity organizationally bottom-up. This can be justified by the intention to satisfy identified needs for measurement (e.g. more detailed information on productivity) as well as the propositions in the literature. Support for this approach was also found in the interviews carried out after the implementation of the measurement systems in the two case services of the first action research study. The approach was operationalized with a method that aggregates productivity change information organizationally bottom-up for purposes of top-level decision-making. In practice, aggregation of information provided by component measures used at the bottom level is carried out organizational level by organizational level arriving at one aggregated index for the whole organization. The method enables the use of various component measures in different services as well as capturing the administrative productivity of each organizational level with respective measures. At the bottom level, productivity should be measured with appropriate methods capturing the unique characteristics of different operations and services. Two component measurement methods for the bottom level measurement, namely weighted unit costs with fixed value for money and productivity matrix, were investigated in this study. The results from interview study II suggest that these methods are suitable for the purposes of operative productivity management in welfare services. Finally, in the process of developing new measurement systems in the cases of this study, the existing knowledge applied in relation to the various phases and tasks needed in performance measurement system development was an essential supporting factor (in addition to the measurement approach and methods chosen) in achieving successful end result.

4 CONCLUSIONS

4.1 Contribution to the existing research

The essential criteria for the contribution of a case research are related to increased knowledge and value from the perspective of the scientific community (Gummesson, 2000, p. 187). The new knowledge provided by this research is related especially to the operative and managerial focus in the examination of productivity and productivity measurement of public services. Much of the existing knowledge on the topic is focused more on the macro level perspective. There may be various purposes in productivity analysis and measurement. The existing research on individual public service sectors (especially health care) is not explicitly linked to management at various organizational levels. The contribution of this research is next examined more closely.

The main contribution of this research relates to the *description of a bottom-up measurement approach supporting productivity management* in the context of large public organizations. The choice of this approach was motivated by the existing literature suggesting that complex phenomena (such as productivity in large public service organizations) may be better captured by splitting them into individual components. It was also noted that this approach may facilitate the commitment of employees and middle management and also increase the relevance of measurement. In addition, prior experiences in the case organization and the results from interview study I (identified needs for measurement) supported this decision. No prior studies describing empirical application of similar approach were found in the literature.

The experiences of the empirical application of the approach were positive. It has potential to solve many of the existing problems related to productivity measurement and management in public organizations as indicated by the results of Article VI. The more detailed examination of this approach was related to two issues: a method aggregating productivity information organizationally bottom-up and operative level (component) measurement.

The contribution of the technical measurement method is related especially to its flexibility and pragmatic principle. With certain limitations different component measures can be used in different services. In addition, the productivity of administration may be measured with its own measures. At the same time, the method enables the linking of measurement carried out at different organizational levels and also the compression of information. In Article IV the new method was compared to a more conventional method with the conclusion that it appears to be at least as valid, reliable and practical but has more relevance.

There are various ways of measuring productivity at the operative level of public organizations. One option is the use of separate productivity-related measures as a part of a balanced performance measurement system. In the case services of this study, a matrix method for multi-dimensional measurement supporting productivity management was used. Consequently, indirect productivity measures were used – not just measures examining the relationship between output/input. The

matrix method was first presented in the literature in the 1980s. However, there is a lack of published studies about it, especially in the context of the public sector. One of the contributions of this study is therefore the empirical investigation of the design and implementation of matrix-based measurement systems in the context of public welfare services.

According to the results from interview study II (reported in Article VI), the matrix method seems to be an efficient tool for operative productivity management in large public organizations. Matrix measurement is practical and satisfies many of the needs identified for measurement: it enables capturing of quality-related aspects, has a 'built-in' target-orientation (information on productivity level regarded as good, average, poor), improves the usage of existing measurement information and provides compact and comprehensible information. It also facilitates benchmarking of similar units which has been regarded essential in modernizing public management.

In addition to the main contribution, two separate contributions are examined in more detail. First, this study provides new knowledge on productivity phenomenon in the context of public services. One of the challenges in the discussion of productivity in general is the ambiguous interpretations of it. Based on the conceptual analysis in Section 1.2.1, a framework of productivity related concepts was presented. It was also argued that there is no specific need to change the common output/input definition in the context of public services. However, in the operationalization of the concept in order to understand the complex phenomenon of productivity comprehensively, it must be split into individual factors (productivity drivers and output elements). This inevitably leads to a broader examination.

Public sector performance models also include other aspects than those related to the examination of services and service production. There is a lack of studies specifically describing productivity phenomenon in public services. Specific productivity-related aspects such as outputs have not been explicitly described in public sector performance models. As a contribution to the prior literature, this study presented a new model for the comprehensive examination of factors affecting productivity (Article I). The model presented can be used for analyzing productivity at both the top and bottom levels of public organizations. In comparison to the existing literature, new aspects of the model are related, e.g. to the more detailed examination of service quality, the role of the customer as well as other 'external factors' in productivity. The addition of the customer aspect in the examination of productivity in public services was also supported by the statistical study in which it was found that customers may affect productivity in various ways.

Many of the studies on the topic of public sector productivity measurement examine and evaluate different measurement methods or the results yielded by the methods. The process of developing productivity measurement systems from the starting point of identifying needs for measurement all the way to the implementation of the systems has been less studied. Second and finally, this study contributes to the existing research by describing the process of developing (design and implementation) measurement systems in a specific setting. Much of the existing research on this topic has been carried out in the private sector with the more general objective of balanced performance measurement. In the specific setting of this study, two key factors affecting the success

of the development process were identified: commitment of the operative level and suitability of the measurement system to the requirements of the organization.

4.2 Contribution to practice

Since an essential part of this research was based on case research carried out as action research, it is necessary to examine the practical contribution of the results (Gummesson, 2000, p. 10; Kasanen et al., 1993; Stam, 2007; Voss et al. 2002). A general issue regarding the practical contribution of this research is the outline and analysis of multidisciplinary and often rather theoretical knowledge about the topic. Due to the practical orientation of this research the theoretical contribution discussed in the previous section may quite easily be turned into a practical contribution.

First, this research provides information on the factors important from the perspective of productivity improvement. Even though some of the more detailed factors may be highly case-specific, the models and classifications described and applied in this study can be utilized in analyzing and prioritizing them. A proper understanding of factors affecting productivity may enable the implementation of enduring practices for productivity improvement instead of harsh ones (e.g. reducing a certain percent of resources systematically in every service). As a simple example, the utilization of service providing capacity could be facilitated by new practices. In child day care the parents could be obligated to give notification of absences of children known in advance. Such days of absences that are known well before may be anticipated by proper human resource management practices. In addition, certain systematic variations in demand (e.g. less demand in Monday and Friday) could be better modeled in order to improve the use of employee resources.

Second, this study mapped a quite versatile list of different ways of utilizing productivity measurement as well as other criteria for measurement. Organizations can use this information in developing measurement systems better suitable for managerial needs based on timely information perceived in the practical setting of a large public organization. The results of this study also illustrate the various needs to improve existing measurement practices.

Third, this research provides knowledge about various ways for improving performance and productivity measurement in public organizations with the emphasis on the operative level. It illustrates the practical application of a measurement approach in which productivity information is gathered organizationally bottom-up. This research also describes the design and implementation of matrix measurement as a component productivity measure in public organizations. Based on the positive experiences and the extent of the implementation in the case organization, it can be argued that matrix is a very promising choice for productivity and performance measurement in any similar public organization.

The fourth practical contribution is related to knowledge about the process of designing and implementing measurement systems in the public sector which can support the measurement-related development projects in organizations. This knowledge is related to the phases and tasks needed as well as various factors supporting and hampering the development process in different phases. The case descriptions can be regarded as illustrative examples since they follow many already known

guidelines for developing measurement systems quite well and also have a relatively successful outcome. An example of such a simple ‘best practice’ is that measurement system development projects must be ambitious but also realistic. There are no perfect measurement systems but there are many systems that reasonably satisfy the managerial needs identified. The deficiencies of measurement systems have to be identified but they cannot be a hindrance for reaching the end in the development work.

4.3 Assessment of the research

4.3.1 Relevance

The first of the four criteria used to assess this research is its relevance, which is also related to the contribution from the perspectives of research and practice (cf. Gummesson, 2000, p. 187). Relevance has been argued to be one of the advantages of case research (Meredith, 1998). The relevance of this research is justified from four viewpoints in this section. First, productivity is an important success factor in every organization. Productivity in public services is a specifically complex and ambiguous phenomenon (cf. Linna et al., 2010). A detailed understanding of the issue both at the top and operative levels of large public organizations is crucial in order to develop appropriate productivity measurement and management systems.

Second, public service productivity measurement is problematic for many reasons, such as output definition. Many challenges have been identified in the literature. However, there is a lack of reported solutions with managerial linkage. This research describes a fresh approach supported by identified needs for measurement, empirical testing, evaluation by the users of measurement as well as some perceived positive impacts. Third, productivity measurement has often been examined at the macro level to analyze the productivity of industries and nations. It has been stated that measurement of productivity as well as performance in general is a highly relevant topic at the operative level of large organizations (Laine, 2005; Ukko, 2009, p. 73). Even though there is potential for productivity improvement in the operative operations of large public organizations, productivity measurement has been rarely linked explicitly to operative level.

The fourth and essential factor improving the relevance of this research is its connection to a long and intense development project with high practical relevance (cf. Westbrook, 1995). This provides a natural setting for examination and access to reality which can be regarded as essential in management research (Gummesson, 2000, p. 14). In light of these arguments, this research can be regarded as relevant enough for a doctoral dissertation.

4.3.2 Validity

The validity of a study refers to the extent to which it covers the phenomenon that researcher aimed to study (Gummesson, 2000, p. 187; McKinnon, 1986). This section examines the validity of this research from the perspective of the two main research questions. The first main research question was related to productivity as a phenomenon which was first investigated at the top level with interview study I. The respondents were chosen by the steering group (with a significant representation from the organization studied) of the development project to gain as comprehensive a

view as possible. The interviewees represented most of the key public services provided by municipal public organizations and can be characterized as major and experienced figures in their own respective fields. The interviewees certainly improved the validity of the results.

The questions for interview study I as a whole were not directly derived from the literature. However, they were loosely connected to the issues of measurement system development (e.g. definition of measurement object and needs for measurement). Another factor affecting the design of the interview was the practical need to gain support for the development work of productivity measurement. The author designed the questions with the help of another researcher. In addition, the questions were evaluated and approved by the steering group of the development work which should have had a positive impact on validity. A factor possibly improving the validity of the results of interview study I, in general, is that the questions were provided to the respondents beforehand. The question related to productivity phenomenon was not restricted with any classifications before the interviews. It is obvious that the respondents had different understandings of productivity. Since productivity is an ambiguous phenomenon, this may be justified by the need to understand genuine conceptions of it. However, the validity of the results may be criticized by questioning whether they describe productivity phenomenon or more broadly performance. On the other hand, argument around concepts can be regarded as irrelevant since there is eventually a need to identify factors that are important from the perspective of improving the efficiency and effectiveness of operations. The analysis of the results from the interview study I were mainly based on the literature and carried out by the author. The validity of the results analyzed is improved by the fact that they were reviewed by another researcher and the representatives of the case organization (cf. Westbrook, 1995; Yin, 2009, p. 41).

The view of the operative level related to productivity phenomenon was examined along with the work of developing measurement systems for individual services. The rationale supporting the selection of the case services was already described in Section 2.4.1. The validity of the results from the perspective of the first main research question is impaired by two factors. First, only 5 services with rather similar characteristics were examined. It is clear that they do not represent anything near the total spectrum of public services but give a reasonable picture of welfare services. However, more comprehensive examination would have required much more work due to the wide range of different services and it was not even the purpose of this study. Second, the ultimate goal of developing measurement systems in the workshops (instead of deep analysis of productivity phenomenon) may have affected the results. However, as demonstrated by the results, a fairly comprehensive list of factors affecting productivity was identified. Many other factors than those measured are represented. Hence, the validity of the results is at a reasonable level.

Productivity impacts were studied at the operative level with the measurement data from a specific public service – child day care. It is clear that the validity of the results leaves a lot to be desired from the perspective of public services in general. However, they are more valid with regard to ‘high-touch’ welfare services. This was also the view of the representatives of other case services (than child day care) when discussing the results in the workshops. Similar studies in other services would certainly have improved the validity of the results. However, this was not possible for practical reasons – there were too few units for statistical analysis in the other case services. The

measures used to provide quantitative data for the analysis were not perfect, which should be taken into account in the assessment of validity. On the other hand, they represented the perspectives of factors affecting productivity fairly comprehensively. Another issue related to validity is the analysis methods used, namely correlations and analysis of variance, which can be regarded as rather simplistic statistical methods. They lack validity from the perspective of a profound understanding related, e.g. to cause-effect relationships. For example, in the results of the correlation analysis the direction is unknown and some other factors than those studied may affect the results (cf. Yin, 2009, p. 42). In addition, no simultaneous impact of multiple factors was studied. In any case, the representatives of the case services studied were able to link many of the results to real life which is an indicator of a reasonable level of validity. Furthermore, the key purpose of the statistical study was to support measurement system development, in which it succeeded well.

The second main research question was related to measurement supporting productivity management. The managerial needs for productivity measurement were first identified at the top and operative levels. Top level needs were investigated in interview study I with five separate questions. Due to the representation from various public departments the results should describe the needs from the top level perspective fairly well. There was a general question as well as more detailed ones related, e.g. to productivity comparisons. Thus, different aspects of the issue should be fairly well represented. A potential challenge in more detailed questions is that they affect the thinking of a respondent. However, the specific issues were considered interesting from the perspective of the literature and the experiences of productivity measurement in the organization studied. The author used data-based classification of the qualitative data from responses which can be justified when exploring a previously unknown issue. As discussed earlier, the review by another researcher as well as the representatives of the case organization improves the validity of the results from interview study I.

The needs for productivity measurement at the operative level were identified in the workshop events of action research study I. Since they were discussed along with the actual work of developing measurement systems, they can be claimed to be realistic and well considered. The validity of the results is impaired by the limited view from only two case services. However, many of the perceptions are in line with the results from interview study I. The needs formulated on the basis of hand-written notes were presented in the workshops and approved by the representatives of the case services.

The research question regarding satisfying managerial needs with measures was also examined organizationally at the top and bottom levels. Both action research studies, used mainly to investigate the research issue in question, were reported in detail in Articles II, IV and V, which facilitates the validity review of external examiner. The validity of both action research studies should be improved by the utilization of an existing structure for the process of developing measurement systems including specified tasks (Bourne, 2003; Lönnqvist, 2004; Neely et al., 2000). The author was the only researcher in the majority of the workshop events related to development. However, there was a constant possibility to correct possible misunderstandings and erroneous assumptions during the many workshops in close interaction between author and the

personnel of the case organization (cf. Westbrook, 1995). The case descriptions were also reviewed and confirmed by representatives of the organization studied, which should improve their credibility.

Since the research question is formulated in a normative manner, it is also appropriate to assess the extent of the testing of the results of development work. The measurement method for aggregating component measurement information was tested in only two case services. The evaluation of the method was carried out in two meetings of the project steering group. This evaluation was interpreted and extended by the author with the support of another researcher in Article IV. Due to limited tests and evaluation by the users of measurement, validity could be improved but it is still at a reasonable level. The measurement method for component measurement (productivity matrix) was tested more extensively and implemented in actual managerial use. In addition, it was evaluated by the representatives of the case services in interview study II reported in Article VI. Consequently, this study can be claimed to provide a valid answer to the research question at the operative level.

Interview study II was also utilized in answering the final research question regarding the factors affecting success in measurement-related development work. The validity of the results is fairly good within the general case-related limitations of this study. The respondents represent different level decision-making in the two case services of action research study I. The choice of respondents was practical – they were personnel who had been involved in the development project and were most familiar with the issues investigated at the time of the interview. The author formulated the questions with the help of another researcher on the basis of the literature, and qualitative content analysis was also carried out based on theory. The author analyzed the results in co-operation with another researcher. In addition, the results were reviewed by the representatives of the case organization.

In general, the validity of this research is at an appropriate level: the research focused on those issues that it was intended to. With regard to both main research questions, the validity of the research is enhanced by examination of both top and bottom organizational levels. This provides a more comprehensive picture of the research issue. As a possible weakness it can be claimed that operative level productivity examination has more emphasis than upper organizational levels. In addition, the operative level examination is limited to five public services. However, this is typical for any case research due to practical limitations. Finally, the research procedures and methods are described in detail enabling validity evaluation by an external evaluator (cf. Gummesson, 2000, p. 186; Yin, 2009, p. 45).

4.3.3 Reliability

The reliability of a study relates to demonstrating that the operations of a study can be repeated with the same results by another researcher (Gummesson, 2000, p. 185; Yin, 2009, p. 40). Reliability is also related to whether the data used can be relied on (McKinnon, 1986). Case studies have been criticized for a lack of reliability (Gummesson, 2000, p. 88). In action research it is unlikely that exactly the same results can be achieved by some other researcher since the intervention of a researcher affects the results (Lönqvist, 2004, p. 238). In addition, changing conditions in

organizations make it difficult to exactly replicate the results of action research. Furthermore, the reliability of action research may be impaired by subjectivity. (Westbrook, 1995) Personal biases may affect what a researcher sees, hears and records (McKinnon, 1986; Voss et al, 2002). Despite the challenges of satisfying the reliability criterion due to the nature of this study, reliability is examined next with specific viewpoints related to the methods used.

In general, several methods and data sources supporting each other were used in the study. For example, statistical methods gave support to the qualitative data gathered during the action research. Linkages between the results from the interviews and action research studies can also be identified. The participation of several researchers, an intense interaction between the researchers and the representatives of the case services should also have a positive effect on reliability (cf. Westbrook, 1995). These factors should reduce the role of the author's own subjectivity in the interpretations.

The reliability of the results from the interviews can always be criticized due to possible variations in respondents' interpretations of questions (Yin, 2009, p. 102). The purpose of the interviews used in this study was mostly explorative with no intention to restrict the responses. The number of interviewees was reasonable in both interview studies, which should improve reliability (cf. Eisenhardt and Graebner, 2007). This is also reflected in the results, in which responses of a similar nature can be identified indicating satisfactory saturation. For example, over half of the respondents of interview study I regarded payment by results as an application of productivity measurement. In the analysis of the interview data, however, interpretation by researchers was needed in analyzing the responses, which inevitably has an effect on the reliability of the results.

The action research was carried out in 5 different case services. Experiences from all the cases were of a similar nature which also supported their simultaneous examination. This also reflects the satisfactory level of the reliability of the perceptions. However, it is impossible to be entirely certain whether additional cases would have provided new information. The reliability of the action research could have been impaired by two factors. First, the author may have been too dominant in the workshop events heavily affecting the results. However, this was not the case, as the results from Article VI demonstrate. Even though the author acted as a facilitator of the development work, the decision-making was democratic. Second, the perceptions and written notes could have been distorted by the author's other experiences (from interviews, other cases etc.). However, the intense interaction between the researchers and the representatives of organization studied in all phases of the research work should have reduced this risk related to reliability of the results.

The reliability of the results obtained through statistical study should be fairly good. The number of organizational units (around 270) examined was reasonably high. The research data used was based mainly on "hard" and objective measures. The data from surveys (e.g. customer satisfaction) can also be regarded as reasonably reliable due to the omission of results from such units in which they were based on only a few responses. Overall the reliability of the internal measurement information should be fairly good since it is widely available in the organization (not only in individual units). Results based on statistical methods are objective, which improves reliability. The statistical analyses were mainly carried out with the measurement information for 2007. Adding other years might have improved the reliability of the results. Finally, conclusions based on the analysis of the

results may naturally have been affected by the researchers' interpretations. However, discussion on conclusions with the representatives of case organization should have improved the reliability.

Overall, it can be stated that the reliability of this study is at a satisfactory level within the limitations of any qualitative case study with intervention by researchers. The long and intense interaction between the researcher and the organization studied should have had a positive impact on the overall reliability of the results (cf. McKinnon, 1986).

4.3.4 Generalization

The fourth and final criterion used to evaluate this research relates to generalization (sometimes referred to as external validity), which relates to the areas to which the result may apply (Gummesson, 2000, p. 187). Case studies have been criticized for limited possibilities for the generalization of findings (Gummesson, 2000, p. 88; Westbrook, 1995). However, the purpose of case studies cannot be in the generalization of results to large populations but to provide detailed understanding of the research issue in specific environments (Meredith, 1998). The main findings of case studies are often applicable at least to some extent in other organizations with similar context ("contextual generalization", see Lukka and Kasanen, 1995).

First, the generalization of the results of this research may be evaluated from the perspective of public organizations in Finland. Even though the City of Helsinki has certain specific characteristics such as size and regional features, there are similar organizations in Finland. More specifically, it may be argued that issues related to top-level examination in both the main research questions have considerable potential for generalization in large public organizations. The generalization of operative level results should be evaluated more carefully. They may be applied in similar welfare services but generalizations to other types of public services such as transportation should be made with caution. This is especially the case with the first main research question. With regard to the other main research question, the potential for generalization should be better. For example, the applicability and potential of matrix-based measurement should be fairly good, at least in such public services in which output definition is challenging. The applicability of the results in small public organizations should be evaluated with care. Some of the issues are possibly not relevant, since the management of the whole organization can be supported by a single measurement system. More research is needed in order to better understand the applicability of the results in such organizations. Finally, the findings related to the factors supporting the development of measurement systems (research question 2C) should be fairly appropriate to the whole public sector due to their general nature.

Second, this research was carried out in Finland, but it should have potential for generalization in the public sectors of other western countries. There are clearly similar challenges and needs for measurement in other countries, as indicated by the literature. Since many of the articles of this study have been published or accepted for publication in international scientific publications, it may be argued that the results also have relevance in other countries than Finland. Factors that should be taken into account in generalizing the findings are at least the differences in organizational structures, legislation and managerial cultures.

Third, generalization can be examined from the perspective of private organizations providing similar services – privatized public services. Organizations providing privatized public services are often smaller in comparison to the empirical context examined in this study. This may have some effect on the applicability of the results. The investigation of productivity phenomenon at the operative level should also be useful in the context of similar privatized public services. In addition, the matrix method for productivity measurement was applied in privatized elderly services as a part of the development project related to this study. According to those experiences, some technical difficulties occurred but no significant challenges or specific reasons why the measurement could not be carried out similarly as in this study. As discussed in Section 1.4.3, however, there may also be other potential methods for productivity measurement due to the availability of price information.

Fourth, generalization of the results in the private sector in general may be examined. At least some of the general issues related to measurement may have potential regardless of the industry. The presented bottom-up approach for productivity measurement may be examined in the context of any large multi-product firms as a potential means of producing overall indices of essential performance factors. The matrix method can be regarded as a way to enhance the use of measurement information in any large organization. Finally, the discussion related to the role of the operative level in supporting successful measurement system development initiatives should also be applicable more widely.

4.4 Suggestions for further research

This study provided new knowledge about the topic of measuring and managing public service productivity. However, there remain many issues requiring more research. Four main paths for future research are presented in this section. They are related both to the limitations of this study and to the experiences from the empirical part of this study.

First, the aggregating measurement method presented in Article IV requires more testing and experiences from different empirical contexts. It has potential to be widely applicable in similar organizations but this can be verified only with further experiences. For example, the purchaser-provider type of organization is clearly a different environment from that of the present study which should be examined in further research. Some adjustments are inevitably needed in further application of bottom-up productivity measurement. Similarly, component productivity measurement, especially with the matrix method, needs more research. The services examined in this study were fairly similar – different public services may pose different challenges. In addition, the applicability of productivity matrix as a key component measure for bottom-up productivity measurement of large public organizations requires more empirical experiences. Finally, this study examined the design and implementation of measurement systems but the actual use of systems as well as impacts and benefits from such use require further research.

Drucker (1999, p. 155), along with many others, has stated that the improvement of knowledge work productivity is a key challenge for the developed countries in the 21st century. Productivity measurement of knowledge and administrative work, which was briefly discussed in Article IV is a

second topic for further research. The issue was also on the agenda in the development project related to this dissertation. Traditionally, knowledge and administrative work has been treated as an input or even ignored in productivity measurement. The bottom-up measurement approach enables a straightforward addition of administrative and knowledge work if the respective component measures are available. This is necessary in the proper application of the approach in order to achieve its full potential. There are two main reasons for this. First, even though changes in administrative and knowledge work productivity have no significant impact on the productivity of large public organizations as a whole, they may have enormous potential for productivity improvement. Second, measurement of knowledge and administrative work productivity has a role in committing personnel by sending the message that the productivity of all the work is under scrutiny – not only that of operative workers. Productivity measurement of administrative and knowledge work is often regarded as a challenging and almost impossible task due to difficulties in output definition and measurement (cf. Okkonen, 2004). This is caused especially by the variation in work content – it is very difficult to identify standard outputs. However, there are also non-standard features in the provision of many case services of this study (e.g. elderly care). Therefore, it would be worth studying whether some of the measurement tools and methods applied, such as the productivity matrix, are applicable in a knowledge work context. In addition, some parts of administrative work are rather standard in nature (e.g. parts of personnel administration) with the potential to apply more conventional productivity measures.

The third topic for further research relates to the productivity measurement of interdepartmental and intradepartmental processes. For example, Kald and Nilsson (2000) have stated that there is a need to improve measurement in order to support process orientation. There is not much research on ‘horizontal’ performance measurement in the public sector (Johnsen, 2005). This issue was also mentioned by several respondents in interview study I. It is essential to maximize the productivity and effectiveness of a municipal organization as a whole. There is no sense in optimizing the productivity of a certain department or departmental section at the expense of another. Work carried out by a certain department may also affect the demand for services in other departments. Due to the extensive use of functional organizational structures in public organizations, it is not possible to systematically apply process-based productivity measurement. However, separate key customer-processes could be identified and measured. In this setting, it would be meaningful to examine productivity from the perspectives of both provider (e.g. resources needed) and customer (e.g. time needed in order to end up with the desired end result).

The fourth theme is the utilization of existing measurement information. According to the representatives of the case services of this study, more efficient utilization of existing measurement information was one of the benefits of applying the matrix method in productivity measurement. This topic could be studied further with at least two approaches. First, statistical methods could be applied in order to learn more from existing measurement information on different public services. It seems that there are not many studies that have utilized operative level measurement information of focused public services (such as child day care in this study). There may be a substantial amount of information that has not been efficiently utilized. Systematic analysis of this information may provide more understanding about the practical means to improve productivity. In macro level analysis, factors affecting productivity may remain too distant from the perspective of operative

management. In article III the impacts of various factors on the productivity of welfare services were studied with correlation and analysis of variance. In further research, more comprehensive measurement data representing a longer time period would be useful. In addition, more sophisticated analysis methods such as regression and cluster analysis, partial correlations as well as DEA analysis could yield even more interesting results from the perspectives of research and practice. One of the specific topics needing more research and also discussed in this research, relates to the role of customers in the productivity of public services.

Another topic that could support the utilization of measurement information of public organizations relates to the improvement of existing measurement and management systems. In the study reported in Article III it was shown that many of the measures available do not properly act as parts of existing measurement systems. Instead, they may be described as separate statistical figures which are, at the worst, not commonly known among personnel.

REFERENCES

- Accounts Commission (1988), *Auditing Guideline – Value-for-Money*, Accounts Commission, Edinburgh.
- Adam, A., Javeri, N., Lessard, M., Lion, J. P., Simard, M., Sorbo, S. (2005), “A case study analysis of the impact of satisfaction and organizational citizenship on productivity”, *Management Research News*, Vol. 28, No. 5, pp. 1–26.
- Allender, H. D. (1997), “Awakening the Power of the Objectives Matrix”, *Industrial Management*, Vol. 39, No. 6, pp. 23–25.
- Ammons, D. N. (2004), “Productivity Barriers in the Public Sector”, in Holzer, M., Lee S. H. (eds.) *Public Productivity Handbook*, Second Edition, Marcel Dekker, New York, NY.
- Ammons, D. N., Coe, C., Lombardo, M. (2001), “Performance-Comparison Projects in Local Government: Participants’ Perspectives”, *Public Administration Review*, Vol. 61, No. 1, pp. 100–115.
- Andriessen, D. (2004), “Reconciling the rigor-relevance dilemma in intellectual capital research”, *The Learning Organization*, Vol. 11, No. 4/5, pp. 393 - 401.
- Antikainen, R. (2006), *Asiantuntijatyön tuottavuusanalyysi - kokemuksia subjektiivisen mittausmenetelmän käytöstä*, Tampereen teknillinen yliopisto, Tuotantotalouden osaston tutkimusraportti 2006:1. (in Finnish)
- Argyris, C., Schön, D. (1978), *Organizational learning: A theory of action perspective*, Reading, Mass, Addison Wesley.
- Atkinson, A. A., Waterhouse, J. H., Wells, R. B. (1997), “A Stakeholder Approach to Strategic Performance Measurement”, *Sloan Management Review*, Vol. 38, No. 3, pp. 25–37.
- Ballentine, J., Brignall, S., Model, S. (1998), “Performance measurement and management in public health services: a comparison of U.K. and Swedish practice”, *Management Accounting Research*, Vol. 9, No. 1, pp 71–94.
- Banker, R. D., Datar, S. M., Kaplan, R. S. (1989), “Productivity Measurement and Management Accounting”, *Journal of Accounting, Auditing and Finance*, Vol. 4, No. 4, pp. 528–554.
- Berry, L., L., Bendapudi, N. (2007), “Health care: a fertile field for service research”, *Journal of Service Research*, Vol. 10, No. 2, pp. 111–122.
- Bierbusse, P., Siesfeld, T. (1997), “Measures that matter”, *Journal of Strategic Performance Measurement*, Vol. 1, No. 2, pp. 6–11.
- Bourne, M. (2003), “The design, implementation and use of performance measurement systems”, *Proceedings of the 3rd International Workshop on Performance Measurement*, Bergamo, 19 – 20. June, pp. 13–22.

- Bourne, M., Franco-Santos, M., Kennerley, M., Martinez, V. (2005), "Reflections on the role, use and benefits of corporate performance measurement in the UK", *Measuring Business Excellence*, Vol. 9, No. 3.
- Bourne, M., Mills, J., Wilcox, M., Neely, A. and Platts, K. (2000), "Designing, Implementing and Updating Performance Measurement Systems", *International Journal of Operations & Production Management*, Vol. 20, No. 7, pp. 754–771.
- Bourne, M., Neely, A., Mills, J., Platts, K. (2003), "Implementing Performance Measurement Systems: a Literature Review", *International Journal of Business Performance Management*, Vol. 5, No. 1, pp. 1–24.
- Bourne, M., Neely, A., Platts, K., Mills, J. (2002), "The success and failure of performance measurement initiatives – Perceptions of participating managers", *International Journal of Operations & Production Management*, Vol. 22, No. 11, pp. 1288–1310.
- Boyle, R. (2006), *Measuring public sector productivity: lessons from international experience*, CPRM Discussion paper 35, Institute of Public Administration, Dublin, Ireland.
- Boyne, G. A. (2002), "Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales", *Public Money & Management*, Vol. 22, No. 2, pp. 17–24.
- Boyne, G. A. (2003), "What is public service improvement?", *Public Administration*, Vol. 81, No. 2, pp. 211–218.
- Brax, S. A. (2007), *Palvelut ja tuottavuus*, Helsinki, Tekes, Teknoliakatsaus 204/2007. (in Finnish)
- Bryman, A., Bell, E. (2007), *Business research methods*, Second edition, Oxford, Oxford University Press.
- Bunge, M. (1973), "On Confusing 'Measure' with 'Measurement' in the Methodology of Behavioral Science" in Bunge, M. (ed.), *The Methodological Unity of Science*, Dordrecht, Holland: Reidel.
- Carson, E., Ranzijn, R., Winefield, A., Marsden, H. (2004), "Intellectual Capital – Mapping Employee Work and Work Group Attributes", *Journal of Intellectual Capital*, Vol. 5, No. 3, pp. 443–463.
- Cavalluzzo, K. S., Ittner, C. D. (2004), "Implementing performance measurement innovations: evidence from government", *Accounting, Organizations and Society*, Vol. 29, No. 3–4, pp. 243–267.
- Chan, Y.C.L. (2004), "Performance measurement and adoption of balanced scorecards – A survey of municipal governments in the USA and Canada", *International Journal of Public Sector Management*, Vol. 17, No. 3, pp. 204–221.
- Chatzkel, J. (2004), "Moving through Crossroads", *Journal of Intellectual Capital*, Vol. 5, No. 2, pp. 337–339.
- Chew, W. (1988), "No-nonsense guide to measuring productivity", *Harvard Business Review*, Vol. 66, No. 1, pp. 110–118.

- City of Helsinki (2002), *This is Helsinki*, <http://www.hel2.fi/kkansl/julkaisut/HelsENG.pdf>, visited 7.4.2008.
- City of Helsinki (2008), *Organizational Chart*, http://www.hel2.fi/ajankohtaista/chart_english.pdf, visited 2.3.2010.
- Coe, C. (1999), "Local Government Benchmarking: Lessons from Two Major Efforts", *Public Administration Review*, Vol. 59, No. 2, pp. 110–123.
- Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J., Battese, G. E. (2005), *An Introduction to Efficiency and Productivity Analysis*, Second Edition, Springer US.
- Collier, P. M. (2006), "Costing Police Services: The Politicization of Accounting", *Critical Perspectives on Accounting*, Vol. 17, pp. 57–86.
- Coughlan, P., Coughlan, D. (2002), "Action research for operations management", *International Journal of Operations & Production Management*, Vol. 22, No. 2, pp. 220–240.
- Craig, C., Harris, C. (1973), "Total Productivity Measurement at the Firm Level", *Sloan Management Review*, Vol. 14, No. 3, pp. 13–29.
- Delaney, J., Huselid, M. (1996), "The impact of human resource management practices on perceptions of organizational performance", *Academy of Management Journal*, Vol. 39, No. 4, pp. 949–969.
- Dervitsiotis, K. N. (1995), "The objectives matrix as a facilitating framework for quality assessment and improvement in education", *Total Quality Management*, Vol. 6, No. 5/6, pp. 563–570.
- Drucker, P. (1963), "Managing for Business Effectiveness", *Harvard Business Review*, Vol. 41, No. 3, pp. 53–60.
- Drucker, P. (1999), *Johtamisen haasteet*, Juva, WSOY. (translated into Finnish from the book *Management Challenges for the 21st Century*)
- Edvinsson, L., Malone, M. S. (1997), *Intellectual Capital: Realizing Your Company's True Value by Finding Its Hidden Brainpower*, HarperBusiness, New York.
- Edwards, D., Thomas, J. C. (2005), "Developing a Municipal Performance-Measurement System: Reflections on the Atlanta Dashboard", *Public Administration Review*, Vol. 65, No. 3, pp. 369–376.
- Eisenhardt, K. M. (1989), "Building theories from case study research", *Academy of Management Review*, Vol. 14, No. 4, pp. 532–550.
- Eisenhardt, K., M., Graebner, M. E. (2007), "Theory Building from Cases: Opportunities and Challenges", *Academy of Management Journal*, Vol. 50, No. 1, pp. 25–32.
- Elinkeinoelämän keskusliitto (2005), *Palvelut 2020 – Kohti palvelujen tulevaisuutta*, Väliraportti, Helsinki. (in Finnish)
- Emory, C. (1985), *Business Research Methods*, 3rd Edition, The Irwin Series in Information and Decision Sciences, Homewood, Illinois.

- Faucett, A., Kleiner, B. H. (1994), "New Developments in Performance Measures of Public Programmes", *International Journal of Public Sector Management*, Vol. 7, No. 3, pp. 63–70.
- Fletcher, J., Snee, H. (1985), "The Need for Output Measurement in the Service Industries: A Comment", *Service Industries Journal*, Vol. 5, No. 1, pp. 73–78.
- Flipo, J-P (1988), "On the Intangibility of Services", *The Service Industries Journal*, Vol. 8, No. 3, pp. 286–293.
- Florentin, E., Omachonu, V. (1991), "Machine centre productivity measurement under conditions of discrete production", *International Journal of Technology Management*, Vol. 6, No.1/2, pp. 131–141.
- Folz, D. H. (2004), "Service Quality and Benchmarking the Performance of Municipal Services", *Public Administration Review*, Vol. 64, No. 2, pp. 209–220.
- Gadrey, J. (1988), "Rethinking Output in Services", *The Service Industries Journal*, Vol. 8, No. 1, pp. 67–76.
- Gaster, L. (1996), "Quality services in local government: a bottom-up approach", *Journal of Management Development*, Vol. 15, No. 2, pp. 80–96.
- Greatbanks, R., Tapp, D. (2007), "The impact of balanced scorecards in a public sector environment – Empirical evidence from Dunedin City Council, New Zealand", *International Journal of Operations & Production Management*, Vol. 27, No. 8, pp. 846–873.
- Green, G., Kennedy, P., McGown, A. (2002), "Management of multi-method engineering design research: a case study", *Journal of Engineering and Technology Management*, Vol. 19, No. 2, pp. 131–140.
- 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.
- Grieves, J., Mathews, B.P. (1997), "Healthcare and the learning service", *The Learning Organization*, Vol. 4, No. 3, pp. 88–98.
- Griffith, J. (2001), "Do Satisfied Employees Satisfy Customers? Support-Services Staff Morale and Satisfaction Among Public School Administrators, Students, and Parents", *Journal of Applied Social Psychology*, Vol. 31, No. 8, pp. 1627–1658.
- Grönroos, C. (2000), *Service Management and Marketing: A Customer Relationship Management Approach*, Chichester, John Wiley & Sons.
- Grönroos, C. (2001), "A Service Quality Model and its Marketing Implications", *European Journal of Marketing*, Vol. 18, No. 4, pp. 36–44.
- Grönroos, C., Ojasalo, K., (2004), "Service productivity: Towards a conceptualization of the transformation of inputs into economic results in services", *Journal of Business Research*, Vol. 57, No. 4, pp. 414–423.
- Gummesson, E. (1998), "Productivity, quality and relationship marketing in service operations", *International Journal of Contemporary Hospitality Management*, Vol. 10, No. 1, pp. 4–15.

- Gummeson, E. (2000), *Qualitative Methods in Management Research*, Second Edition. Sage Publications, Thousand Oaks (CA).
- Gupta, A. (1995), "Productivity measurement in service operations: a case study from the health-care environment", *Managing Service Quality*, Vol. 5, No. 5, pp. 31–35.
- Hackman, S. T. (2008), *Production Economics - Integrating the Microeconomic and Engineering Perspectives*, Springer Berlin.
- Halinen, I., Korhonen, S. (2008), *Parasta tuottavuutta – Kuntatuottavuus -hankkeen loppuraportti*, Suomen kuntaliitto, Helsinki. (in Finnish)
- Hannula, M. (1999), "Expedient Total Productivity Measurement", *Acta Polytechnica Scandinavica, Industrial Management and Business Administration Series*, No. 1
- Hannula, M. (2002), "Total productivity measurement based on partial productivity ratios", *International Journal of Production Economics*, Vol. 78, No. 1, pp. 57–67.
- Hannula, M., Lönnqvist, A. (2002), *Suorituskyvyn mittauksen käsitteet*, Metalliteollisuuden Kustannus, Helsinki. (in Finnish)
- Hautakangas, S. and Heikkinen, J. (2008), "Miten tuottavuuden kehitystä mitataan julkisissa palveluissa?", *Tieto&Trendit 8/2008*, Tilastokeskus. (in Finnish)
- Hautakangas, S., Heikkinen, J., Laine, S., Seppänen, O. (2007), "Julkisten palvelujen tuottavuuden mittaaminen Tilastokeskuksessa", in Kangasharju, A. (ed.), *Hyvinvointipalvelujen tuottavuus – Tuloksia opintien varrelta*, VATT Publications 46, Helsinki, pp. 139–157. (in Finnish)
- Hill, T. P. (1977), "On goods and Services", *Review of Income and Wealth*, Vol. 123, No. 4, pp. 315–338.
- Hirsjärvi, S., Hurme, H. (2004), *Tutkimushaastattelu – teemahaastattelun teoria ja käytäntö*, Helsinki University Press, Helsinki. (in Finnish)
- Hjerppe, R., Kangasharju, A. (2003), "Miksi tuottavuutta olisi lisättävä?", in Hjerppe, R., Kangasharju, A., Vuorento, R. (eds.) (2003), *Kunnalliset palvelut – Terveysten- ja vanhustenhuollon tuottavuus*, Helsinki, VATT-julkaisuja 37, pp. 11–14. (in Finnish)
- Hodgkinson, A. (1999), "Productivity measurement and enterprise bargaining – the local government perspective", *International Journal of Public Sector Management*, Vol. 12, No. 6, pp. 470–481.
- Hollingsworth, B. (2003), "Non-parametric and parametric applications measuring efficiency in health care", *Health Care Management Science*, Vol. 6, No. 4, pp. 203–218.
- Hollingsworth, B. (2008), "The measurement of efficiency and productivity of health care delivery", *Health Economics*, Vol. 17, No. 10, pp. 1107–1128.
- Hood, C. (1995), "The New Public Management in the 1980s: Variations on a Theme", *Accounting, Organizations and Society*, Vol. 20, No. ½, pp. 93–109.
- Hovila, H., Okkonen, J. (2006), *Kokemus organisaation voimavaraksi*, eBRC Research Reports 32, Tampere.

- Huff, L., Fornell, C., Anderson, E. W. (1996), "Quality and productivity: contradictory and complementary", *Quality Management Journal*, Vol. 4, No. 1, pp. 22–39.
- Hui, C., Lam, S.S.K., Schaubroeck, J. (2001), "Can good citizens lead the in providing quality service? A field quasi experiment", *Academy of Management Journal*, Vol. 44, No. 5, 988–995.
- Hupli, M., Kaarna, T., Kauppinen, R., Kärri, T. (2006), *Sairaalan ohjausmallit: potilas kannattavaksi*, Tutkimusraportti 173, Lappeenranta teknillinen yliopisto. (in Finnish)
- Huxham, C., Vangen, S. (2003), Researching organizational practice through action research: case studies and design choices, *Organizational Research Methods*, Vol. 6, No. 3, pp. 383–403.
- Häkkinen, U. (2008), "Palveluketjut: tapa mitata tuotosta ja tehostaa tuotantoa", in Ilmakunnas, S. (ed.), *Hyvinvointipalveluja entistä tehokkaammin – uudistusten mahdollisuuksia ja keinoja*, VATT publications 48, Helsinki, pp. 97–114. (in Finnish)
- Ilmakunnas, S. (2008), "Hoidon ja hoivan tuotannon taloudelliset haasteet", in Ilmakunnas, S. (ed.), *Hyvinvointipalveluja entistä tehokkaammin – uudistusten mahdollisuuksia ja keinoja*, VATT publications 48, Helsinki, pp. 1–16. (in Finnish)
- Institute of Industrial Engineers (1990), *Industrial Engineering Terminology, revised edition*, Industrial Engineering and Management Press, Norcross, Georgia.
- Jacobs, R. (2001), "Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis", *Health Care Management Science*, Vol. 4, No. 2, pp. 103–115.
- Jarrar, Y., Schiuma, G. (2007), "Measuring performance in the public sector: challenges and trends", *Measuring Business Excellence*, Vol. 11, No. 4, pp. 4–8.
- Johnsen, Å (2005), "What Does 25 Years of Experience Tell Us About the State of Performance Measurement in Public Policy and Management?", *Public Money & Management*, Vol. 25, No. 1, pp. 9–17.
- Johnsen, Å, Norreklit, H., Vakkuri, J. (2006), "Introducing a Nordic Perspective on Public Sector Performance Measurement", *Financial Accountability & Management*, Vol. 22, No. 3, pp. 207–212.
- Johnston, R., Clark, G. (2008), *Service Operations Management: Improving Service Delivery*, 3rd Edition, Pearson Education Limited.
- Johnston, R., Jones, P. (2004), "Service productivity. Towards understanding the relationship between operational and customer productivity", *International Journal of Productivity and Performance Management*, Vol. 53, No. 33, pp. 201–213.
- Jääskeläinen, A., Kujansivu, P., Väisänen, J. (2010), "Productivity Analysis of Public Services – An Application of Data Mining", in Syväjärvi, A, Stenvall, J. (eds.) *Handbook of Research on Data Mining in Public and Private Sectors: Organizational and Government Applications*, IGI Global publisher, pp. 83–104.
- Jääskeläinen, A., Lönnqvist, A. (2008), "Public Service Productivity: How to Capture the Intangible Elements?", In: O'Sullivan, K. (ed.) *5th International Conference on Intellectual Capital*,

Knowledge Management & Organisational Learning (ICICKM-2008), 9-10 October, 2008, New York, USA, pp. 233–240.

- Jääskeläinen, A., Lönnqvist, A. (2009), “Designing Operative Productivity Measures in Public Services”, *VINE - Journal of Information and Knowledge Management Systems*, Vol. 39, No. 1, pp. 55–67.
- Kald, M., Nilsson, F. (2000), “Performance measurement at Nordic Companies”, *European Management Journal*, Vol. 18, No. 1, pp. 113–127.
- Kangasharju, A. (2008), “Tuottavuus osana tuloksellisuutta”, in Ilmakunnas, S. (ed.), *Hyvinvointipalveluja entistä tehokkaammin – uudistusten mahdollisuuksia ja keinoja*, VATT Publications 48, Helsinki, pp. 169–202. (in Finnish)
- Kangasharju, A., Kirjavainen, T., Luoma, K., Rätty, T. (2007), “Koulutuspalvelujen tuottavuuden ja tehokkuuden mittaaminen”, in Kangasharju, A. (ed.) *Hyvinvointipalvelujen tuottavuus – Tuloksia opintien varrelta*, VATT Publications 46, Helsinki, pp. 119–137. (in Finnish)
- Kaplan, R. S. (2001), “Strategic Performance Measurement and Management in Nonprofit Organizations”, *Nonprofit Management & Leadership*, Vol. 11, No. 3, pp. 353–370.
- Kaplan, R. S., Atkinson, A. A. (1998), *Advanced Management Accounting*, third edition, Prentice Hall, Upper Saddle River.
- Kaplan, R. S., Norton, D. P. (1992), “The Balanced Scorecard – Measures That Drive Performance”, *Harvard Business Review*, Vol. 70, No. 1, pp. 71–79.
- Kaplan, R. S., Norton, D. P. (1996), *The Balanced Scorecard – Translating Strategy into Action*, Harvard Business School Press, Boston, Massachusetts
- Kaplan, R. S., Norton, D. P. (2001), *The Strategy-Focused Organization. How Balanced Scorecard Companies Thrive in the New Business Environment*, Harvard Business School Press, Boston.
- Kaplan, R. S., Norton, D. P. (2004), *Strategy Maps. Converting Intangible Assets into Tangible Outcomes*, Boston, Harvard Business School Publishing Corporation.
- Kasanen, E., Lukka, K., Siitonen, A. (1993), “The Constructive Approach in Management Accounting Research”, *Journal of Management Accounting Research*, Vol. 5, pp. 243–264.
- Kaydos W. (1999), *Operational Performance Measurement: Increasing Total Productivity*, St. Lucie Press, Florida.
- Klassen, K. J., Russell, R. M., Chrisman, J. J. (1998), “Efficiency and Productivity Measures for High Contact Measure”, *Service Industries Journal*, Vol. 18, No. 4, pp. 1–18.
- Kloot, L. and Martin, J. (2000), “Strategic Performance Management: A Balanced Approach to Performance Management Issues in Local Government”, *Management Accounting Research*, Vol. 11, No. 2, pp. 231–251.
- Kransdorff, A. (1995), “Exit interviews as an induction tool”, *Management Development Review*, Vol. 8, No. 2, pp. 37–40.
- Kujala, J., Lillrank, P., Kronström, V., Peltokorpi, A. (2006), “Time-based management of patient processes”, *Journal of Health Organization and Management*, Vol. 20, No. 6, pp. 512–524.

- Kujansivu, P. (2008), *Intellectual Capital Management – Understanding Why Finnish Companies Do Not Apply Intellectual Capital Management Models*, Tampere University of Technology, Publication 747.
- Käpylä, J., Jääskeläinen, A., Lönnqvist, A. (2010), "Identifying future challenges for productivity research: evidence from Finland", *International Journal of Productivity and Performance Management*, Vol. 59, No. 7, pp. 607–623.
- Labro, E., Tuomela, T.S. (2003), "On bringing more action into management accounting research: process considerations based on two constructive case studies", *European Accounting Review*, Vol. 12, No. 3, pp. 409–442.
- Laine, J. (2005), "Tuottavuuden arviointi osana sosiaalitaloudellista tutkimusta – Näkökulmia lasten päivähoiton tuottavuustutkimukseen", *Yhteiskuntapolitiikka*, Vol. 70, No. 6, pp. 659–666. (in Finnish)
- Lancer Julnes, P., Holzer, M. (2001), "Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation", *Public Administration Review*, Vol. 61, No. 6, pp. 693–708.
- Larsson, S. O (1995), "Quality Adjustment of Productivity Analysis in the Public Sector" in Halachmi, A., Bouckaert, G. (Eds.) *Public Productivity Through Quality and Strategic Management*, IOS Press, pp. 39–51.
- Lawton, A., McKevitt, D. and Millar, M. (2000), "Coping with Ambiguity: Reconciling External Legitimacy and Organizational Implementation in Performance Measurement", *Public Money & Management*, Vol. 20, No. 3, pp. 13–19.
- Lehtonen, V.-M. (2007), *Henkilöstöjohtamisen tehostaminen valtionhallinnossa henkilöstötilinpäätösinformaation avulla – empiirinen tutkimus Suomen valtionhallinnossa tuotettavan henkilöstötilinpäätösinformaation arvosta johtamisessa*, Helsinki, Svenska handelshögskolan. (in Finnish)
- Likierman, A. (1993), "Performance indicators: 20 early lessons from managerial use", *Public Money & Management*, Vol. 13, No. 4, pp. 15–22.
- Linna, P., Pekkola, S., Ukko, J., Melkas, H. (2010), "Defining and measuring productivity in the public sector: managerial perceptions", *International Journal of Public Sector Management*, Vol. 23, No. 3, pp. 300–320.
- Loggerenberg van, B., Cucchiaro, S. (1982), "Productivity Measurement and the Bottom Line", *National Productivity Review*, Vol. 1, No. 1, pp. 87–99.
- Lorenzo, J. M. P., Sanchez, I. M. G. (2007), "Efficiency evaluation in municipal services: an application to the street lighting service in Spain", *Journal of Productivity Analysis*, Vol. 27, No. 3, pp. 149–162.
- Lovelock, C. H. (1983), "Classifying Service to Gain Strategic Marketing Insights", *Journal of Marketing*, Vol. 47, No. 3, pp. 12–17.
- Lukka, J. (2000) "The key issues of applying the constructive approach to field research" in Reponen, T. (ed.) *Management Expertise for the New Millennium: In Commemoration of the*

- 50th Anniversary of the Turku School of Economics and Business Administration. Publications of Turku School of Economics and Business Administration, Series A-1:2000, pp. 113–128.
- Lukka, K., Kasanen, E. (1995), “The Problem of Generalizability: Anecdotes and Evidence in Accounting Research”, *Accounting, Auditing & Accountability Journal*, Vol. 8, No. 5, pp. 71–90.
- Lumijärvi, I. (1999), *Tasapainotetun mittariston malli ja kunta-alan tuloksellisuusarviointi*, Helsinki, Työturvallisuuskeskus. (in Finnish)
- Lupton, T. (1991), “Organizational Change: Top-down or bottom-up Management?”, *Personnel Review*, Vol. 20, No. 3, pp. 4–10.
- Lynch, R. L., Cross, K. F. (1995), *Measure Up! How to Measure Corporate Performance*, Blackwell Publishers Inc, Malden.
- Lönnqvist, A. (2002), *Suorituskyvyn mittauksen käyttö suomalaisissa yrityksissä*, Licenciate Thesis, Tampereen teknillinen korkeakoulu, tuotantotalouden osasto, Tampere. (in Finnish)
- Lönnqvist, A. (2004), *Measurement of Intangible Success Factors: Case Studies on the Design, Implementation and Use of Measures*, Tampere, Tampere University of Technology, Publication 475.
- Maliranta, M. (2007), “Mitä tuottavuus on ja miten mitata sitä?” in Maliranta, M., Yläanttila, P. (eds.) *Kilpailu, innovaatio ja tuottavuus*, Helsinki, ETLA, pp. 29–48. (in Finnish)
- Mammone, J. (1980), “Productivity Measurement: A Conceptual Overview”, *Management Accounting*, Vol. 61, No. 12, pp. 36–43.
- Matta, K. (1989), “A Goal-oriented Productivity Index for Manufacturing Systems”, *International Journal of Operations & Production Management*, Vol. 9, No. 4, pp. 66–76.
- McAdam, R. Hazlett, S.-A. and Casey, C. (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.
- McCunn, P. (1998), “The Balanced Scorecard: the eleventh commandment”, *Management Accounting*, Vol. 76, No. 11, pp. 34–36.
- McKinnon, J. (1986), “Reliability and Validity in Field Research: Some Strategies and Tactics”, *Accounting, Auditing and Accountability*, Vol. 1, No. 1, pp. 34–54.
- McLaughlin, C. P., Coffey, S. (1990), “Measuring Productivity in Services”, *International Journal of Service Industry Management*, Vol. 1, No. 1, pp. 46–64.
- Meredith, J. (1998), “Building operations management theory through case and field research”, *Journal of Operations Management*, Vol. 16, No. 4, pp. 441–454.
- Midwinter, A. (1994), “Developing Performance Indicators for Local Government: The Scottish Experience”, *Public Money & Management*, Vol. 14, No. 2, pp. 37–43.
- Miles, M. B., Huberman, A. M. (1994), *Qualitative Data Analysis: an expanded sourcebook*, Second Edition, Thousand Oaks, Sage.

- Miller, D. (1984), "Profitability = productivity + price recovery", *Harvard Business Review*, Vol. 62, No. 3, pp. 145–153.
- Miller, R. T., Murnane, R. J., Willett, J. B. (2008), "Do worker absences affect productivity? The case of teachers", *International Labour Review*, Vol. 147, No. 1, pp. 71–89.
- Ministry of Finance (2007), *Valtiovarainministeriön hallinnonalan tuottavuuden mittaamishankkeen loppuraportti*, Valtiovarainministeriön julkaisuja 1/2007, Helsinki. (in Finnish)
- Misterek, S., Dooley, K., Anderson, J. (1992), "Productivity as a performance measure", *International Journal of Operations & Production Management*, Vol. 12, No. 1, pp. 29–45.
- Modell, S. (2001), "Performance measurement and institutional processes: a study of managerial responses to public sector reform", *Management Accounting Research*, Vol. 12, No. 4, pp. 437–464.
- Neely, A., Adams, C. (2000), *Perspectives on Performance: The Performance Prism*, Centre of Business Performance, Cranfield Business School, UK.
- Neely, S., Gregory, M. and Platts, K. (1995), "Performance measurement system design: a literature review and research agenda", *International Journal of Operations & Production Management*, Vol. 15, No. 4, pp. 80–116.
- Neely, A., Mills, J., Gregory, M., Richards, H., Platts, K., Bourne, M. (1996), *Getting the Measure of Your Business*, Findlay, London.
- Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., Kennerley, M. (2000), "Performance Measurement System Design: Developing and Testing a Process-Based Approach", *International Journal of Operations & Production Management*, Vol. 20, No. 10, pp. 1119–1145.
- Niemi, M. (1997), "Data envelopment -analyysi", in Uusi-Rauva, E. (ed.) *Tuottavuus – mittaa ja menesty*, Kauppakaari, Helsinki, pp. 263–270. (in Finnish)
- Nordgren, L. (2009), "Value creation in health care services – developing service productivity", *International Journal of Public Sector Management*, Vol. 22, No. 2, pp. 114–127.
- Norreklit, H., (2000), "The balance on the Balanced Scorecard – a critical analysis of some of its assumptions", *Management Accounting Research*, Vol. 11, No. 1, pp. 65–88.
- Näsi, S., Hokkanen, A., Latvanen, J. (2001), *Johdon laskentatoimen hyväksikäyttö ja kehittämistarpeet erilaisissa kunnallisissa toimintayksiköissä*. Jyväskylän yliopiston taloustieteiden julkaisuja N:o 125/2001. (in Finnish)
- Ojasalo, K. (1999), *Conceptualizing Productivity in Services*, Helsingfors, Svenska handelshögskolan, 75.
- Ojasalo, K. (2003), "Customer Influence on Service Productivity", *S.A.M. Advanced Management Journal*, Vol. 68, No. 3, pp. 14–19.

- Okkonen, J. (2004), *The use of performance measurement in knowledge work context*, Doctoral Dissertation, e-Business Research Center Research reports 9, Tampere, Tampere University of Technology.
- Olkkonen, T. (1994), *Johdatus teollisuustalouden tutkimustyöhön*, Toinen painos, Report no. 152, Teknillinen korkeakoulu, Otaniemi.
- Parasuraman, A. (2002), "Service quality and productivity: a synergistic perspective", *Managing Service Quality*, Vol. 12, No. 1, pp. 6–9.
- Parkkinen, P. (2007), *Väestön ikääntymisen vaikutukset kuntatalouteen*, VATT publications 136, Helsinki. (in Finnish)
- Penrose, E. (1995), *The theory of the growth of the firm*, Oxford, Oxford University Press.
- Phelps, M. (2009), "Total public service output and productivity", *Economic & Labour Market Review*, Vol. 3, No. 8 pp. 45–55.
- 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, T. H. (2003), *Measuring performance in public and nonprofit organizations*, Jossey-Bass, U.S.
- Poister, T. H., Streib, G. (1999), "Performance Measurement in Municipal government: Assessing the State of the Practise", *Public Administration Review*, Vol. 59, No. 4, pp. 325–335.
- Pollanen, R. (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. Summa, H. (1997), "Trajectories of Reform: Public Management Change in Four Countries", *Public Money & Management*, January-March, pp. 7–18.
- Pritchard, R. D. (1995), *Productivity Measurement and Improvement: Organizational Case Studies*, Praeger Publishers, Westport, CT.
- Radnor, Z., McGuire, M. (2004), "Performance management in the public sector: fact or fiction?", *International Journal of Productivity and Performance Management*, Vol. 53, No. 3, pp. 245–260.
- Rajala, T., Meklin, P. (2008), "Laskentatoimi organisatorisen tehottomuuden tunnistajana", in Hyvönen, T., Laine, M., Mäkelä, H. (eds.) *Laskenta-ajattelun tutkija ja kehittäjä: Professori Salme Näsi 60 vuotta*, Tampereen yliopisto, Taloustieteiden laitos, pp. 97–110. (in Finnish)
- Rantanen, H. (1995), *The effects of productivity on profitability: a case study at firm level using an activity-based costing approach*, Lappeenranta, Lappeenranta University of Technology 45.
- Rantanen, H., Holtari, J. (1999), *Yrityksen suorituskyvyn analysointi*, Lappeenranta, Lappeenrannan teknillinen korkeakoulu, Tuotantotalouden osasto, Tutkimusraportti 112. (in Finnish)
- Rantanen, H., Kulmala, H. I., Lönnqvist, A., Kujansivu, P. (2007), "Performance Measurement Systems in the Finnish Public Sector", *International Journal of Public Sector Management*, Vol. 20, No. 5, pp. 415–433.

- Rautiainen, A. (2004), *Kuntien suorituksen mittaus ja tuloksellisuuskäsitteistö: arviointi ja kehittäminen*, Jyväskylän yliopisto taloustieteiden tiedekunta, julkaisuja 140/2004.
- Regan, W. J. (1963), "The Service Revolution", *Journal of Marketing*, Vol. 27, No. 3, pp. 57–62.
- Reicheld, F. F., Sasser, W. E. (1990), "Zero defections: quality comes to services", *Harvard Business Review*, Vol. 68, No. 5, pp. 105–111.
- Riggs, J. L. (1981), *Production Systems: Planning, Analysis and Control*, 3rd Edition, New York, Wiley.
- Riggs, J. L. (1986), "Monitoring with a matrix that motivates", *Performance Management Magazine*, Vol. 4, No. 3, pp. 13–19.
- Rosen, E. D. (1993), *Improving Public Sector Productivity – Concepts and Practice*, Sage, Newbury Park, Calif.
- Ruggiero, J. (1996), "Efficiency of educational production: an analysis of New York school districts", *Review of Economics and Statistics*, Vol. 78, No. 3, pp. 499–509.
- Saari, S. (2000), *Tuottavuuden mittaus osana kannattavuuden mittausta teollisuusyrityksessä: Mittausmenetelmien vertaileva tutkimus*, Tampere University of Technology, Publications 290. (in Finnish)
- Sahay, B. S., (2005) "Multi-Factor Productivity Measurement Model for Service Organization", *International Journal of Productivity and Performance Measurement*, Vol. 54, No. 1, pp. 7–22.
- Salminen, M. and Viitala, M.-L. (ed.) (2006) *Handbook on Performance Management, Governance and Accountability*, 2/2006, Edita Publishing PLC, Helsinki. (Publication available at <http://www.ministryoffinance.fi/publications>).
- Sampson, S. E., Froehle, C. M. (2006), "Foundations and Implications of a Proposed Unified Services Theory", *Production and Operations Management*, Vol. 15, No. 2, pp. 329–343.
- Sanderson, I. (2001) "Performance Management, Evaluation and Learning in Modern Local Government", *Public Administration*, Vol. 79, No. 2, pp. 297–313.
- Schlesinger, L.A., Zornitsky, J. (1991), "Job satisfaction, service capability, and customer satisfaction: an examination of linkages and management implications", *Human Resource Planning*, Vol. 14, No. 2, pp. 141–149.
- Sherwood, M. K. (1994) "Difficulties in the measurement of service outputs", *Monthly Labor Review*, Vol. 117, No. 3, pp. 11–19.
- Silvestro, R., Fitzgerald, L., Johnston, R., Voss, C. (1992), "Towards a Classification of Service Processes", *International Journal of Service Industry Management*, Vol. 3, No. 3, pp. 62–75.
- Simons, R. (2000), *Performance Measurement & Control Systems for Implementing Strategy*, Prentice Hall, New Jersey.
- Simpson, H. (2006), *Productivity in public services*, Working Paper 7/165, Centre for Market and Public Organisation, University of Bristol.

- Simpson, H. (2009), "Productivity in Public Services", *Journal of Economic Surveys*, Vol. 23, No. 2, pp. 250–276.
- Singh, H., Motwani, J., Kumar, A. (2000), "A review and analysis of the state-of-the-art research on productivity measurement", *Industrial Management & Data Systems*, Vol. 100, No. 5, pp. 234–241.
- Sink, D. (1983), "Much Ado About Productivity: Where Do We Go From Here", *Industrial Engineering*, Vol. 15, No. 10, pp. 36–48.
- Sink, D. (1985), *Productivity Management: Planning, Measurement and Evaluation, Control and Improvement*, New York, John Wiley & Sons.
- Slack, N., Chambers, S., Johnston, R (2001), *Operations Management*, 3rd ed., Harlow, Prentice Hall.
- Sosiaalivirasto (2008), *Sosiaaliviraston toimintakertomus 2008*,
[<http://www.hel.fi/wps/wcm/connect/8bdc92004a176c87baf9fa3d8d1d4668/toimintakertomus2008.pdf?MOD=AJPERES>], 15.6.2009. (in Finnish)
- Stam, C. D. (2007), *Knowledge productivity: designing and testing a method to diagnose knowledge productivity and plan for enhancement*, Doctoral thesis, University of Twente.
- Statistics Finland (2006), *Kansantalouden tilinpidon menetelmä- ja aikasarjauudistus aiheuttaa tasokorjauksia aikasarjoihin*, Announcement 31.3.2006,
http://www.stat.fi/ajk/tiedotteet/v2006/tiedote_022_2006-03-31.html, 22.3.2010. (in Finnish)
- Statistics Finland (2008), *Kuntien ja kuntayhtymien tuotoksen volyymin, kokonaispanoksen volyymin ja kokonaistuottavuuden muutos toimialoittain vuosina 2001–2008, prosenttia edellisestä vuodesta*, http://www.stat.fi/til/kktu/2008/kktu_2008_2010-02-05_tau_001_fi.html, 23.3.2010. (in Finnish)
- Stainer, A. (1997), "Capital input and total productivity management", *Management Decision*, Vol. 35, No. 3, pp. 224–232.
- Stainer, A., Stainer, L. (1998) "Strategic change in public services: a productivity and performance perspective", *Strategic Change*, Vol. 7, No. 2, pp. 111–119.
- Sveiby, K.-E. (1997) *The New Organizational Wealth: Managing and Measuring Knowledge-Based Assets*, Berrett-Koehler Publishers Inc, San Francisco.
- Talbot, C. (1999) "Public Performance – towards a new model?", *Public Policy and Administration*, Vol. 14, No. 3, pp. 15–32.
- Tangen, S. (2005), "Demystifying productivity and performance", *International Journal of Productivity and Performance Management*, Vol. 54, No. 1, pp. 34–46.
- Thomas, B.E., Baron, J.P. (1994), *Evaluating knowledge worker productivity: literature review*, Interim Report, No. FF-94/27 [www.cecer.army.mil/kws/tho_lit.htm#abs], 11.3.2009.
- Toivanen, J. (2001) *Implementation of the Balanced Scorecard and the Current Use of the Balanced Scorecard in Finland*, Doctoral dissertation, Lappeenranta University of Technology, Lappeenranta. (In Finnish)

- Tuomi, J., Sarajärvi, A. (2009), *Laadullinen tutkimus ja sisällönanalyysi*, Helsinki, Tammi. (in Finnish)
- Ukko, J. (2009), *Managing Through Measurement: a framework for successful operative level performance measurement*, Acta Universitatis Lappeenrantaensis, 348, Lappeenranta.
- Uusi-Rauva, E. (1997) (ed.), *Tuottavuus – mittaa ja menesty*, 2. Painos, Helsinki, Kauppakaari Oy. (in Finnish)
- Vakkuri, J. (2003), “Research Techniques and Their Use in Managing Non Profit Organizations. An Illustration of DEA Analysis in NPO Environments”, *Financial Accountability & Management*, Vol. 19, No. 3, pp. 243–264.
- Vakkuri, J., Meklin, P. (2006), “Ambiguity in Performance Measurement: A Theoretical Approach to Organisational Uses of Performance Measurement”, *Financial Accountability & Management*, Vol. 22, No. 3, pp. 235–250.
- Van Aken, J. E. (2007) “Design Science and Organization Development Interventions: Aligning Business and Humanistic Values”, *Journal of Applied Behavioral Science*, Vol. 43, No. 1, pp. 67–88.
- Van Helden, J. G. (2005), “Researching Public Sector Transformation: The Role of Management Accounting”, *Financial Accountability & Management*, Vol. 21, No. 1, pp. 99–133.
- Van Helden, G. J., Johnsen, Å, Vakkuri, J. (2008), “Distinctive research patterns on public sector performance measurement of public administration and accounting disciplines”, *Public Management Review*, Vol. 10, No. 5, pp. 641–651.
- Van Looy, B., Gemmel, P., Desmet, S., Van Dierdonck, R., Serneels, S. (1998), “Dealing with productivity and quality indicators in a service environment: some field experiences”, *International Journal of Service Industry Management*, Vol. 9, No. 4, pp. 359–376.
- Vehmanen, P. (2008), “Fair value measurements” in Hyvönen, T., Laine, M., Mäkelä, H. (eds.) *Laskenta-ajattelun tutkija ja kehittäjä: Professori Salme Näsi 60 vuotta*, Tampereen yliopisto, Taloustieteiden laitos, pp. 37–58.
- Viitamo, E. (2007), *Productivity of Business Services – Towards A New Taxonomy: Research report for the project SCR-06*, SC-Research, Lappeenranta University of Technology, Department of Industrial Management.
- Viitamo, E. (2009), *Service Management and Economy – Merging Analysis on Service Productivity*, Lappeenranta University of Technology, Research Report 212, Lappeenranta.
- Voss, C., Tsikriktsis, Frohlich, M. (2002), “Case research in operations management”, *International Journal of Operations & Production Management*, Vol. 22, No. 2, pp. 195–219.
- Vuorinen, I., Järvinen, R., Lehtinen, U. (1998), “Content and measurement of productivity in the service sector. A conceptual analysis with an illustrative case from the insurance business”, *International Journal of Service Industry Management*, Vol. 9, No. 4, pp. 377–396.
- Wakefield, K. L., Blodgett, J. G. (1999), “Customer Response to Intangible and Tangible Service Factors”, *Psychology & Marketing*, Vol. 16, No. 1, pp. 51–68.

- Westbrook, R. (1995), "Action research: a new paradigm for research in production and operations management", *International Journal of Operations & Production Management*, Vol. 15, No. 12, p. 6–20.
- Westlund, A., Löthgren, M. (2001), "The interactions between quality, productivity and economic performance: the case of Swedish pharmacies", *Total Quality Management*, Vol. 12, No. 3, pp. 385–396.
- Williams, S. (1991), "Strategy and Objectives", in Neale, F. (ed.) *The Handbook of Performance Management*, Institute of Personnel Management, Short Run Press Ltd, UK.
- Wisniewski, M., Olafsson, S. (2004), "Developing balanced scorecards in local authorities: a comparison of experience", *International Journal of Productivity and Performance Management*, Vol. 53, No. 7, pp. 602–610.
- Wisniewski, M., Stewart, D. (2004), "Performance measurement for stakeholders – The case of Scottish local authorities", *International Journal of Public Sector Management*, Vol. 17, No. 3, pp. 222–233.
- Xu, K., Jayaram, J., Xu, M. (2006), "The effects of customer contact on conformance quality and productivity in Chinese service firms", *International Journal of Quality & Reliability Management*, Vol. 23, No. 4, pp. 367–389.
- Yin, R. K. (2009), *Case Study Research – Design and Methods*, Fourth Edition, Thousand Oaks (Calif.), Sage Publications.

APPENDICES

Appendix 1: Examples of productivity matrices in two case services

Appendix 1: Examples of productivity matrices in two case services

	Unit cost for weighted care day, €	Customer satisfaction of parents (survey result 1-5)	Number of children per number of employees	Degree of utilization of premises %	Percentage of S2 children, %	Rate of short sickness absences of employees, %
Measurement result						
10	35.00	4,40	5,7	100	40	2,0
9	35.50	4,37	5,6	99	36	2,5
8	36.00	4,34	5,4	98	32	3,0
7	37.00	4,31	5,2	97	28	3,5
6	38.00	4,28	5,0	96	24	4,0
5	39.00	4,24	4,8	94	20	5,0
4	40.00	4,20	4,6	92	16	6,0
3	42.00	4,10	4,4	90	12	7,0
2	45.00	4,00	4,2	85	8	8,0
1	48.00	3,80	3,9	80	4	9,0
0	52.00	3,60	3,6	75	0	10,0
Score						
Weight	45	15	15	15	5	5
Weighted score						
Total score						

Productivity matrix in child day care services

	Unit cost for weighted care day, €	Customer satisfaction of family members (survey result 1-5)	Customer satisfaction of handicapped people (percentage of positive answers based on a survey)	Percentage of utilization of care places, %	Rate of short sickness absences of employees, %
Measurement result					
10	31,00	4,50	95	100	1,5
9	32,00	4,45	94	99	1,7
8	33,00	4,40	92	98	2,0
7	34,00	4,35	90	96	2,5
6	35,00	4,30	88	94	3,0
5	36,00	4,20	86	92	3,5
4	38,00	4,10	84	90	4,0
3	42,00	4,00	81	87	5,0
2	46,00	3,90	78	84	6,0
1	50,00	3,70	75	80	8,0
0	55,00	3,50	70	75	10,0
Score					
Weight	45	15	10	15	15
Weighted score					
Total score					

Productivity matrix in care for mentally handicapped people

PART 2: PUBLICATIONS

ARTICLE I

Jääskeläinen, A. (2010)

Identifying factors affecting public service productivity

International Journal of Services Technology and Management

Vol. 14, No. 4, pp. 360–375

ARTICLE II

Jääskeläinen, A., Lönnqvist, A.

Public Service Productivity: How to Capture Outputs?

International Journal of Public Sector Management (accepted for publication)

ARTICLE III

Jääskeläinen, A., Kujansivu, P., Lönnqvist, A. (2009)
Learning from Existing Performance Information
*Proceedings of 5th Conference on Performance Measurement and
Management Control, September 23-25, 2009, Nice, France*

ARTICLE IV

Jääskeläinen, A., Uusi-Rauva, E.

Bottom-up Approach for Productivity Measurement in Large Public Organizations

International Journal of Productivity and Performance Management

(accepted for publication)

ARTICLE V

Jääskeläinen, A. (2009)

Identifying a Suitable Approach for Measuring and Managing Public Service
Productivity

Electronic Journal of Knowledge Management

Vol. 7, No. 4, pp. 447–458

ARTICLE VI

Jääskeläinen, A., Sillanpää, V. (2010)

Evaluating the factors of success in the performance measurement system
development of two public services

Proceedings of 10th EURAM 2010 Conference, May 19-22, 2010, Rome, Italy