Danish Journal of Management & Business

Vol. 78, no. 3 and 4, November 2014

Formerly published as Ledelse & Erhvervsøkonomi

3 & 4/2014

Editorial

Performance Management

Per Nikolaj Bukh and Jan Mouritsen

Abstract

Performance is an ambiguous concept. In this editorial to the special issue on *performance management* we take the lack of clear definitions as an invitation to explore what is meant by performance, how performance can be perceived and how we can understand managerial activities related to performance. The papers in the issue consider performance management at an organisational level and address new developments and thinking in relation to performance management. The focus of the papers demonstrates how performance management constitutes an umbrella for the study of formal processes that organizations use in attempting to implement their strategic intent, and to adapt to the circumstances in which they operate. Performance management draws on many disciplines although the concept appears to have merged with management accounting and control highlighting that the central role of budgeting as a control technique has declined.

1. Introduction

Organisations in all sectors of the economy attempt to improve performance at a general level but the irony is that as a concept it is ill defined. It is ambiguous and holds different potentialities for different stakeholders (e.g. Dooren *et al.* 2010; Moynihan 2008; Talbot 2010). It can also be defined in relation to a very wide range of entities – such as organisational units, individual employees and managers, lateral processes, value chains, inter-organisational collaborations and societies – where both design and behaviour are important antecedents to the functioning of performance management in all those contexts.

It is not strange that the aim of performance management is to enhance performance. In other words, the problem of performance management is to identify the managerial and organisational efforts that are expected to produce performance effects. Since managers and organizations continually adapt their performance management practices to changes in dynamic environments, it is likely that new ways of performing as well as new ways of managing performance are developed. The openness of the term performance as well as the impossibility of clear general definitions of what is meant both by managing and by performance have been suggested as hindrances for building a cumulative body of research (Baxter &

Chua 2003; Ferreira & Merchant 1992). However, it can also be seen as an invitation to explore in much further detail how instances of performance management can be understood and explained.

The old management adage »You can't manage what you don't measure« has an intuitive appeal to practitioners developing performance management systems. It seems somehow obvious that unless you measure something you do not know if it improves or worsens. If there is no knowledge about improvements in performance variables it is difficult to see how and when interventions can be contemplated (e.g. Cokins 2009; Davenport & Harris 2007; Kaplan & Norton 2008). This view is a key driver in much modern public sector reforms as well as development of strategic performance measurement systems in the private sector.

The problem is that it may not be so clear what should count as performance. As Colin Talbot suggests it may not be possible to quite to determine whether organisations »have a 'real' performance that is 'out there', independent of any cognition or perspective, or is it merely a 'construct' dependent on the observer« (Talbot 2010, p. 54). This means that an appeal to performance is not really a solution; it is a problem in the sense that it requires not only manages but also researchers to make their situated notion of performance open. Even it performance measurement points to the idea that numbers have objectivity, it is also clear that the object that the numbers reflects with some idea of objectivity may not themselves be beyond doubt and requirement of explanation. Measurements are thus »selected, interpreted and used by actors in different ways consistent with their institutional interests« as Moynihan (2008, p. 8) puts it.

In the call for papers for this special issue on performance management the aim is to address new developments and thinking in relation to performance management. This can include new perspectives, new techniques and related analysis designed to improve individual, group organisational and inter-organisational performance. Performance means different things to different people, and the dimensions of performance differ between various situations and contexts. The papers in the special issue all consider performance management at an organisational level. But the performance to be managed emanates from efforts to manage produced at many levels.

2. Narrowing down the performance concept

Talbot (2010) traces theories of organisational performance to the literature on organisational effectiveness after the end of World War II. He points out one line of argument consistent with the literature from the 1980'es and early 1990'es, which aimed at improving organisations as e.g. Peters & Waterman's (1982) famous study, *In search of Excellence*.

Another line of argument related to literature focusing on the managerial activities and can be tracked back to Robert Anthony who in 1965 introduced the management control concept as distinguished from strategic planning and operational

control (Anthony 1965). Management control systems provide in the current meaning »information that is intended to be useful to managers in performing their jobs and to assist organizations in developing and maintaining viable patterns of behaviour (Otley 1999, p. 365). As a common denominator for this literature, the management of improvements of performance stands out.

Performance can, however, be given many and often ambiguous meanings and can be associated with a range of activities. Dubnick (2005) mentions mundane activities such as opening a door or re-enacting a musical and concludes that »performance stands in distinction to mere 'behaviour' in implying some degree of intent« (Dubnick 2005, p. 391). Also David Otley (2001) emphasizes intent in his characterisation of performance management as it provides »an umbrella under which we can study the more formal processes that organizations use in attempting to implement their strategic intent, and to adapt to the circumstances in which they operate« (Otley 2001, p. 250).

In other words, performance is about intentional behaviour. Yet, the question of what to perform is left open. Although the central issue in management control remains in the contemporary understanding of performance management the way management control its expressed has as noted by Otley (2003, p. 315) changed since it was introduced in 1965. Thus, performance management reflects the same issues and concerns as management control, however the central role of budgeting as a control technique has declined.

Many aspects of performance can in commercial firms often be related to financial terms. Sometimes the performance in producing outputs or procuring inputs can be translated to financial number by valuations. In other situations can various non-financial aspects of performance as demonstrated by Nielsen *et al.* (2014) be linked analytically to financial terms.

To some extent, public sector organisations may differ, as their performance cannot be narrowed down to a question of profitability. Rather the public sector is said to be characterised by activities performed by professionals, public services that are multiple-value entities and are developed and delivered in co-production mode (See also Bruijn, 2007). However, performance management systems in the form of principles of financing decentralised units, e.g. schools, hospitals, agencies etc. using case based formula funding (cf. Smith 2007) often structures the public sector very similar private firms as is seen e.g. in Jakobsen's (2014) article on performance budgeting in this issue.

Increasingly, the financing of the public organisations, e.g. upper secondary schools and language schools for Danish immigrants (Dahler-Larsen & Pihl-Thingvad 2014), is dependent on the performance in various dimensions. There are complex relations between public services and financial appropriations. These relations are in need of being discovered and analysed rather than assumed. The problem is probably not that budgets and financial incentives are present; the

issue is much more how different kinds of budgetary mechanism and incentive schemes are related to the production of outputs and outcomes that are understood as socially desirable.

It is probably a good idea to look for the specificities to develop a nuanced theory of how performance management will be a resource for improving performance (Bruijn 2007, Moynihan 2008; Pollitt & Bouckaert 2011). Such an emphasis on various forms of performance management systems has also associated its implementation in many parts of the Danish public sector. Adopting a broader view of performance management may bridge the organisational and institutional aspects of performance with the techniques adopted.

Performance management has been an essential part of many new management accounting and control techniques that has been developed over the past thirty years (cf. Otley 2001, p. 248). Not only performance measurement but also many other management accounting techniques, as e.g. Activity Based Costing, is included in the performance management concept when Otley's lens is adopted. Further, issues within business intelligence (e.g. Cokins 1999) and analytics (e.g. Davenport & Harris 2007; Davenport et al. 2010) are also within the performance management umbrella.

3. The papers in the special issue

The first two articles in the special issue focus on the balanced scorecard (BSC). Since this performance management system was introduced by Kaplan & Norton in the beginning of the 1990'es (Kaplan & Norton 1992; 1996), it has been understood as moving from a performance measurement to a strategic management system (Kaplan & Norton 2001); by the introduction of the Closed-Loop Management model (Kaplan & Norton 2008) it has also evolved into a comprehensive performance management model encompassing a number of separate principles and techniques. BSC is widely used in all parts of the world, including Denmark where the first implementations were seen around 1995.

The balanced scorecard is defined in the books by Robert S. Kaplan and David Norton. However, in a national context it is shaped by supply-side actors, e.g. consulting firms, who adopt the concept according to their own specialties and experiences. In the first article Dag Øivind Madsen (Madsen 2014) examines how the Balanced Scorecard has been interpreted and used in Denmark. Drawing on interviews with suppliers and users of the BSC, Madsen shows how these different actors have attached different kinds of meaning to the BSC concept.

Madsen identifies the two main interpretations in Danish practice: the BSC as a 'performance measurement system' and as a 'strategic management system', the latter representing performance being managed and not merely measured. Interestingly, there is a distance between how consultants and user organizations interpret and attach meaning to the concept. While suppliers interpret and promote the

BSC as a strategic management system, users appear more likely to interpret and use the concept as a performance measurement system.

The second article is authored by Steen Nielsen, Erland H. Nielsen, Anders Jacobsen, Lars Bjørn Pedersen (Nielsen et al. 2014) and focuses on how system dynamics modelling can contribute to an increased understanding of a firms strategy based on Kaplan & Norton's (2001) strategy map. A strategy map represents a high-level model of causal relationships among strategic objectives and Kaplan (2009, p. 1268) has previously suggested that a detailed system dynamics model could incorporate causal linkages that have both time delays and more complex feedback loops than can be represented in a traditional strategy map. However, empirical applications of this technique have been sparse and the article by Nielsen et al. is an important contribution to the literature.

The article demonstrates how it is possible to enhance understanding of the causal relationships in the strategy map by this approach and thus help to develop activities and other decision that connect even better with the company's strategy. By combining the idea of balanced scorecards as a causal loop system with systems thinking, the authors address not only the comprehensiveness of the balanced scorecard, but also the methodological approach which by the use of differential equations makes it possible to identify the strength of its causal relations.

The third article by Peter Dahler-Larsen and Signe Pihl-Thingvad (Dahler-Larsen & Pihl-Thingvad 2014) examines the stress effects performance management. The article is based on a comprehensive survey among three groups of public sector employees: language teachers for Danish emigrants, upper secondary school teachers and job consultants in job centres. The article focuses on differences between these groups and points out a series of specific factors potentially leading to stress in relation performance management for these three types of jobs.

The choice of employee groups is in a Danish performance management context an innovative and interesting setup. Job centres have used performance measurements as incentives to a very high degree in the public sector in Denmark: Employees experience an intense public control in the form of statistical indicators and benchmarking and this area has much public and political attention. This public debate has often argued that the performance management focus here harms the working environment.

Also upper secondary schools and language teachers are, as explained in the article, exposed to performance measurement and performance management, although to a lesser degree than the job centres. Important for the study, the professional self-understanding of upper secondary teachers is based on academic degrees, although their various disciplines range from humanities over social science and business to the natural sciences, while, there is no clear basic profession among job consultants and language teachers.

Contrary to expectations the stress level is found to be lower among the job consultants than in the two other employee groups. In analysing the results, the authors suggests that stress in each group is best explained by a classical model that looks at work load, clarity of indicators, and lack of control over the work situation. The most stressful group experienced freedom in their work, were very engaged in the work, set high standards for themselves, but on the other hand, the most stressed group felt that performance indicators did not help to create clarity about what should be done at work. Even though further analysis of differences between individuals within group could shed more light on the mechanisms discussed in this article, the article points to the importance of role of clarity in performance management and to the role of the employee´s own work standards when performance management systems are designed.

In the fourth article Bente Bjørnholt, Jeppe Agger Nielsen and Andrej Christian Lindholst (Bjørnholt *et al.* 2014) investigate the differences and similarities between performance management systems in two different areas of public services in Denmark: Social services (eldercare) and technical services (park services). As suggested by Otley (2001), see also Moynihan *et al.* (2011), performance management systems should be designed and adapted to achieve intended organizational strategies and the design of a performance management system is contingent on a number of external and internal organizational factors, including the task complexity of particular services.

From this contingency perspective, Bjørnholt *et al.* (2014) suggest that the characteristics of performance management systems should be expected to vary across policy areas in the public sector. Contrary to expectations for divergence due to differences in task complexity, the authors find that the characteristics of performance management systems in the two policy areas tend to converge with respect to purpose, target setting, measurement, and motivational mechanisms.

Bjørnholt *et al.* (2014) trace the reasons for the convergence and suggest a set of complementary propositions in addition to the already established proposition on the influx from government-wide policies, which can assist in explaining convergence in the characteristics of performance management systems in the public sector. Specifically, it is proposed that convergence can occur due to similarities in policy-specific reforms, institutional pressures, and complementarity between political and managerial needs

Finally, the fifth article by Mads Leth Felsager Jakobsen (Jakobsen 2014) is about the degree of adoption of a specific performance management technique, performance budgeting, at Danish hospitals. As frequently documented in the public management literature formal adoption does not necessarily imply real changes (Dooren et al. 2010; Moynihan 2008). Drawing on Olivier's (1991) framework of strategic responses to institutional processes Jakobsen (2014) examines how and why the hospitals as well as specific orthopaedic wards respond to various institutional forces.

Overall, Jakobsen concludes that the pattern of adoption mirrors the international trend of partial adoption of performance management systems. Hospitals and wards have had their reimbursement linked to output, but this has only partly been accompanied by financial autonomy. This partial adoption has been particularly prevalent at the ward level. Furthermore, the patterns of adoption seem to be the combined result of an institutional pressure directed at the formal elements of formula-based performance budgeting and the fit between performance budgeting and management capacity.

In particular, management capacity could be an important explanation for the different patterns of adoption at the hospital and the ward level although it requires more studies and analysis to understand what it requires to design, implement and use technically advanced performance management systems. The performance budgeting concept implemented by the Danish hospitals imply that rather advanced formula funding models (cf. Smith 2007) should be implemented. The question is whether the managerial capacity and the accounting sophistication at the hospitals is at a level where full adoption of performance budgeting is technically feasible at all.

4. Conclusion

Approaching performance management as an umbrella term where we can study the more formal processes that organizations use in attempting to implement their strategic intent, and to adapt to the circumstances in which they operate (cf. Otley 1999, 2001) has the strength that it includes a wide range of performance management issues and management devices and that it recognizes that it is the combination of performance elements that is of interest, rather than the isolated parts of a performance management system.

Performance management remains a rich and interesting field of research with many opportunities for new and creative studies that combines existing fields of research and paves new roads to new practises. When studying the operation of systems as complex as performance management it is clear that increasingly more concerns, entities and objects are taken into account. What appears to be a simple technical interest in setting targets and associated funding rules, performance management is a broad organisational field that ranges from individuals' problems with stress to organisational strategy, and in-between are the many diverse mechanism that allow organisations to operate make objectives, coordinate and delegate on many diverse organisational values.

References

Anthony, R.S. 1965. *Management planning and control systems: A framework for research*. Boston: Harvard Graduate School of Business.

Baxter, J. & W.F. Chua. 2003. Alternative management accounting research: whence and whither. *Accounting, Organizations and Society*, Vol. 28, pp. 97-126

Bjørnholt, B., J.A. Nielsen & A.C. Lindholst. 2014. Three propositions on why characteristics of performance management systems converge across policy areas with different levels of task complexity. *Danish Journal of Management & Business*, this issue.

Bruijn, H.de. 2007. *Managing performance in the public sector*, second edition. London: Routledge.

Cokins, G. 2009. *Performance Management: Integrating strategy, execution, methodologies, risk and analytics*. Hoboken: John Wiley & Sons.

Dahler-Larsen, P. & S. Pihl-Thingvad. 2014. How Performance Indicators can lead to Stress: A group level analysis on performance measurements and stress. *Danish Journal of Management & Business*, this issue.

Davenport, T.H., J.G. Harris & R. Morison. 2010. *Analytics at work: Smarter decisions, better results*. Boston: Harvard Business Press.

Davenport, T.H. & J.G. Harris 2007. *Competing on analytics: The new science of winning*. Boston: Harvard Business School Press.

Dooren, W. van; G. Bouckaert & J. Haligan. 2010. *Performance Management in the Public Sector*. London: Routledge.

Dubnick, M. 2005. Accountability and the promise of performance: In search of mechanisms. *Public Performance & Management Review*, vol 28, pp. 376-417.

Ferreira, L.D. & K.A. Merchant. 1992. Fiels Research in Management Accouting and control: A reviewand evaluation. *Accounting, Organizations and Society,* Vol. 5, No. 4, pp.3-34.

Jakobsen, M.L.F. 2014. The partial adoption of performance budgeting at Danish hospitals. *Danish Journal of Management & Business*, this issue.

Kaplan, R.S. 2009. Conceptual Foundations of the Balanced Scorecard. I *Handbook of Management Accounting Research*, Vol. 3, C.S. Chapman, A.G. Hopwood og M.D. Shields (red.). Amsterdam: Elsevier.

Kaplan, R.S. & D.P. Norton 1992. The balanced scorecard – Measures that drive performance. *Harvard Business Review*, Vol. 70, No. 1, pp. 71-79.

Kaplan, R.S. & D.P. Norton 1996. *The Balanced Scorecard: Translating strategy into action*. Boston: Harvard Business School Press.

Kaplan, R.S. & D.P. Norton 2001. *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston: Harvard Business School Press.

Kaplan, R.S. & D.P. Norton 2008. *The Execution Premium: Linking Strategy to Operations for Competitive Advantage.* Boston: Harvard Business School Press.

Madsen, D.Ø. 2014. Interpretation and use of the Balanced Scorecard in Denmark: Evidence from suppliers and users of the concept. *Danish Journal of Management & Business*, this issue.

Moynihan, D.P. 2008. *The Dynamics of performance management: Constructing information and reform.* Washington, D.C.: Georgetown University Press.

Nielsen, S, E.H. Nielsen, A. Jacobsen & L.B. Pedersen. 2014. Management Accounting and Business Analytics: An example of using System Dynamics Modelling to facilitate the design of a Balanced Scorecard. *Danish Journal of Management & Business*, this issue.

Oliver, C. 1991. Strategic responses to institutional processes. *The Academy of Management Review,* 16(1), 145-179.

Otley, D. 1999. Performance management: a framework for management control systems. *Management Accounting Research*, vol. 10, pp. 363-382.

Otley, D. 2001. Extending the boundaries of management accounting research: Developing systems for performance management. *British Accounting Review*, vol. 33, pp. 243-261.

Otley, D. 2003. Management control and performance management: whence and whiter? *The British Accounting Review*, Vol. 35, pp. 309-326.

Peters, T. & R. Waterman. 1982. *In search of excellence: Lessons from Americas best run companie*s. New York: Harper & Row.

Pollitt, C. & Bouckaert, G. 2011. *Public management reform: A comparative analysis – New public management, governance, and the neo-weberian state*. Oxford University Press.

Smith, P. 2007. *Formula funding of public services*. London: Routledge.

Talbot, C. 2010. *Theories of performance: Organizational and service improcements in the public domain*. Oxford: Oxford University Press.

Interpretation and use of the Balanced Scorecard in Denmark: Evidence from suppliers and users of the concept

Dag Øivind Madsen

Abstract

This paper examines how the Balanced Scorecard (BSC) has been interpreted and used in Denmark. Drawing on interviews with suppliers and users of the BSC, the paper shows how these different actors have interpreted and attached meaning to the BSC concept. The BSC's 'interpretive space' and 'boundary object' characteristics lead to different interpretations and mobilizations of the concept in practice. There is a multiplicity in terms of what the concept means on both the supply side and the demand side. Two main interpretations can be identified in Danish practice: the BSC as (1) a 'performance measurement system'; and (2) as a 'strategic management system'. Interestingly, there is a mismatch between how consultants and user organizations interpret and attach meaning to the concept. While suppliers interpret and promote the BSC as a strategic management system, users appear more likely to interpret and use the concept as a performance measurement system. The paper discusses possible explanations for these differences.

1. Introduction

Since its introduction in 1992, the Balanced Scorecard (BSC) has attracted the interest of managers, consultants and academics (Hoque, 2014). *Harvard Business Review* has referred to Kaplan and Norton's (KN) (1992) BSC concept as one of the most important management ideas of the last 75 years (Sibbet, 1997: 12). Studies show that it is widely used by managers in different parts of the world (Marr, 2005; Rigby & Bilodeau, 2013; Silk, 1998; Speckbacher, Bischof, & Pfeiffer, 2003). At the same time, research indicates that the concept's reception varies across different countries, e.g. in terms of the awareness, adoption and stages of implementation.

Studies show that the trajectory of the BSC concept in a given national context is shaped by supply-side actors (e.g. consulting firms and management gurus) who adapt the concept according to their own professional specialties and experiences (Braam, Benders, & Heusinkveld, 2007; Braam, Heusinkveld, Benders, & Aubel, 2002) or local market preferences (Ax & Bjørnenak, 2005). Similarly, demand-side actors (i.e. organizations and their managers) shape the concept as they interpret

and draw on the concept in different ways in practice (Braam, 2012; Dechow, 2012; Hansen & Mouritsen, 2005; Modell, 2012; Wagensveld, 2013).

1.1 Purpose and contribution

Against the background outlined above, the purpose of the paper is to examine how the BSC concept has been interpreted and mobilized in Denmark. Using data from interviews with those on both the supply side and demand side, I show how different actors interpret and attach meaning to the BSC concept.

The paper engages with the emerging literature which studies the diffusion, translation and popularization of the BSC (e.g. Ax & Bjørnenak, 2005; Braam et al., 2007; Cooper, Ezzamel, & Qu, 2012; Qu, 2004). These studies have employed various theoretical perspectives including 'management fads and fashions', 'actor-network theory' and 'travelling ideas'. The current paper attempts to show how insights from these three perspectives can be used to study how the BSC has been interpreted and used in Denmark. In addition, unlike most previous research on popular management concepts (cf. Benders & Van Veen, 2001; Clark, 2004), the paper relies not only on data from the supply side, but also data about interpretations and use of the concept on the demand side.

The combination of supplier and user data allows for a more 'balanced' view of how the BSC concept is interpreted and used in Denmark. It also means that it is possible to analyze variances in the way different types of actors interpret and attach meaning to the concept. Finally, the paper brings new insight to how the BSC concept is interpreted and used in a Danish business context. Existing literature about the BSC in a Danish context is limited (e.g. Christensen & Bukh, 2013; Hansen & Mouritsen, 2005; Nielsen & Sørensen, 2004); therefore this study should be of interest to academics and practitioners alike.

1.2 Structure

The remainder of the paper is structured as follows. Part two discusses different theoretical perspectives which can be used to study the emergence and evolution of management concepts such as the BSC. In part three, the study's research approach is outlined. In part four and five, I describe the empirical findings about how the BSC is interpreted and used on the supply side and demand side, respectively. Section six follows a discussion of the findings, and the concluding part a discussion of the paper's main shortcoming and suggestions for future studies.

2. Theoretical background

2.1 Management concepts

Management concepts are *»prescriptive, more or less coherent views on management«* (Benders & Verlaar, 2003: 758). Management fashions constitute a subset of the total supply of management concepts and are those *»management concepts that relatively speedily gain large shares in the public management discourse«* (Jung & Kieser, 2012: 329). There are four main characteristics of management

concepts that give them the potential to 'flow' (Røvik, 2002) in the management community, i.e. to become 'fashionable' concepts:

- the concept is labeled in a catchy way, usually using a three-letter acronym (Røvik, 1998);
- performance improvements are promised to potential adopters of the concept (Røvik, 1998; ten Bos, 2000);
- the concept is presented as universally applicable (Strang & Meyer, 1993); and
- there is wide room for interpretation of the concept – so-called 'interpretive space' (Benders & Van Veen, 2001; Clark, 2004; Giroux, 2006).

2.2 The BSC as a management concept

The BSC can be considered a management concept since it exhibits the four characteristics identified above. It is labeled in catchy way, and triggers positive associations to sports and fitness as most managers want to be 'balanced' and know their 'score' relative to their competitors (Kieser, 2002: 176). In KN's conceptual literature about the BSC, it is underscored how the concept can lead to 'excellent' performance, and that the concept can be applied in both the private and the public sector, and across countries and cultures. The last characteristic is arguably particularly important for the BSC. Many researchers have pointed out that the BSC can be interpreted in different ways (Ax & Bjørnenak, 2005; Braam, 2012; Braam et al., 2007; Malmi, 2001; Nørreklit, 2003; Speckbacher et al., 2003), is flexible and malleable, and can stand for many different things in practice (Aidemark, 2001; Dechow, 2012; Hansen & Mouritsen, 2005; Modell, 2004; Modell, 2012).

2.3 The interpretive space of the BSC and its implications

The interpretive space of the BSC means that different interpretations and understandings of the BSC exist, both on the supply side and the demand side. On the supply side, KN have played a leading role in shaping how the concept has evolved, via their best-selling books, conference presentations and their consulting organization (Cooper et al., 2012). They have over time moved from presenting the BSC primarily as a performance measurement system to a strategic management system (Braam & Nijssen, 2004). In addition, other actors (e.g. management gurus and consulting firms) have also thrown their hats in the ring and written and developed their own variations of the concept (e.g. Niven, 2005). Researchers have also found that local suppliers of the BSC have actively shaped the content of the BSC to make it fit better with the local context (Ax & Bjørnenak, 2005). As a whole, these studies show that the BSC is presented in different ways by different supply-side actors.

These variations of the BSC concept can be found not only on the supply side, but also on the demand side. For example, in a study of the BSC in the Netherlands, Braam et al. (2002) distinguished between two main translations of the BSC. In the 'technical' translation the BSC is used primarily in the 'accounting way' as a tool for performance measurement, whereas in the 'organizational' translation the BSC is used more dynamically in the strategy development process. In a related follow-

up study, Braam (2012) refers to the 'interpretative variability' of the BSC when it is used in organizational change projects.

Several empirical studies have classified different types of BSC use in organizational praxis. Speckbacher et al. (2003) distinguishes between three types of BSC use, where type 1 is close to the original formulation of the BSC as a performance measurement system, while type 3 is a more advanced use of the BSC as a strategic management system. In a similar vein, Lawrie and Cobbold (2004) identify three generations of BSC use, where the third generation represents the most ambitious version of the BSC. Brudan (2005) distinguishes between five categories of BSC use, ranging from the BSC as primarily an information system to an advanced version incorporating advanced elements such as strategy maps and strategic destinations.

Taken together, these studies show that the BSC lends itself to different interpretations and translations, and that many variations of the BSCs can be found on the supply side and in use in organizational praxis. This means that the concept is likely to have a unique trajectory in a given local context since different actors (both suppliers and users) interpret and mobilize the concept in different ways. The next section describes three theoretical approaches that can offer insight into how the BSC emerges and evolves in a given context.

2.4 Three theoretical views of the BSC

Different theoretical perspectives have been used to study diffusion, popularization and translation of the BSC, most notably management fashion theory (MFT), actor-network theory (ANT), and the Scandinavian institutional theory (SIT). These theoretical approaches offer three views of the nature and characteristics of the BSC concept:

- the BSC as a 'management fashion' (MFT);
- the BSC as a 'boundary object' (ANT); and
- the BSC as a 'travelling idea' (SIT).

2.4.1 The BSC as a 'management fashion'

The first view is that the BSC can be considered a 'management fashion'. This research stream has drawn on MFT (e.g. Abrahamson, 1996; Benders & Van Veen, 2001; Carson, Lanier, Carson, & Guidry, 2000) to explain the emergence and evolution of the BSC concept. Primarily these studies have explored how the BSC has evolved at the macro level, i.e. the national level. For example, Malmi (2001) studied the adoption and use of the BSC in Finnish organizations, and found that management fads and fashions were among the main motives driving BSC adoption. In a study of the BSC in Sweden, Ax and Bjørnenak (2005) pointed out the important role played by fashion-setters such as consultants and conference organizers in the early phase of BSC diffusion in Sweden. In the Netherlands, researchers have found that fashion-setters have played a role as providers of BSC-related discourse in the business media (Braam et al., 2007; Braam et al., 2002).

The focus of the research carried out using an MFT approach has been on the concept's macro-level evolution, i.e. patterns in the dissemination and diffusion of the BSC concept. Less attention has been paid to how the concept has been interpreted at the micro level, i.e. by individual organizations. However, even though management fashion researchers have focused on macro-level dissemination and diffusion patters, they have also pointed out the flexible nature of the BSC, and that the concept can be interpreted and customized in different ways on both the supply side (Ax & Bjørnenak, 2005) and the demand side (Braam, 2012; Modell, 2009). Ax and Bjørnenak (2005) argue that fashion-setters make such adaptations as part of the fashion-setting process to appeal to the local market. The interpretive flexibility of the BSC may also explain its widespread diffusion and 'fashion potential', as the BSC is perceived as potentially useful and appealing to a wide range of actors in different contexts.

2.4.2 The BSC as a 'boundary object'

The second research stream has studied the popularization of the BSC using insights and inspiration from ANT (Briers & Chua, 2001; Latour, 1987), a theory which is frequently used in management accounting studies (Justesen & Mouritsen, 2011). In particular, Star and Griesemer's (1989) notion of a 'boundary object' has been used in previous BSC studies. In the words of Star and Griesemer (1989: 393) *»boundary objects are objects which are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites.«* The BSC can be considered a boundary object since the concept has considerable interpretive space and is flexible, but also has certain stable core characteristics that can be identified across different interpretations, such as the four perspectives and a combination of leading (forward-looking) and lagging (backward-looking) indicators (Hansen & Mouritsen, 2005).

Hence, it can be seen that the ANT-inspired research stream views the emergence and evolution of the BSC from a slightly different vantage point compared to MFT-inspired researchers. ANT-inspired researchers pay comparatively less attention to the role of fashion-setters on the supply side. However, there are exceptions. For example, Qu (2004) studied the rise and dissemination of the BSC, with particular emphasis on how KN have built a network and enrolled allies around the concept. In a recent study of the popularization of BSC, Cooper et al. (2012) show how KN have popularized the concept via their consulting organizations, books, conference presentations, etc. Hence, these two ANT-inspired studies have some commonalities with studies carried out using a MFT lens. For example, they emphasize the role of supply-side actors such as KN in the diffusion and popularization process. In particular they focus on how KN built networks of support around the concept to facilitate and drive diffusion and institutionalization of the BSC.

Other studies in this research stream focus more on how the concept has been enacted on the demand side. According to ANT-inspired researchers, it is primarily

the BSC's boundary object characteristics that give the concept its 'fashion potential' since organizations can draw on the BSC in a variety of ways to address different organization-specific problems. For example, Hansen and Mouritsen (2005) focus on how the BSC is used as a boundary object in organizational praxis. In their four Danish case organizations they find that the BSC is plastic enough to stand for different things. Hansen and Mouritsen (2005: 144) note that the BSC is capable of being *»...bent towards purposes so that a local identity could be upheld and yet... that the BSC [would be] implemented, rather than something else.«* Furthermore, they find that the BSC is interpreted differently in the four case organizations and mobilized in relation to each organization's specific problems, e.g. business planning, process reengineering, or benchmarking.

2.4.3 The BSC as a 'travelling idea'

The third view is that the BSC can be considered a 'travelling idea'. This SIT-inspired research stream focuses on how global management ideas such as the BSC 'travel' can become contextualized in local praxis (Czarniawska & Sevón, 2005; Czarniawska & Sevòn, 1996; Sahlin-Andersson & Engwall, 2002; Sahlin & Wedlin, 2008). Specifically, the focus of this line of BSC research is on how the concept attains a local flavor as it circulates and 'travels' between different contexts. Most of the studies have been conducted at the micro level and have explored how concepts are implemented and translated on the user side of the market. Although not the main focus of the theory, this research stream acknowledges that suppliers act as 'carriers' (Sahlin-Andersson & Engwall, 2002) and 'circulators' (Powell, Gammal, & Simard, 2005) who shape how users interpret and use concepts.

Several researchers have studied the BSC as a 'travelling idea'. For example, Wongkaew (2007) studied how the BSC was translated in a Thai financial services organization. In a similar vein, Nilsen (2007) studied the translation of the BSC idea in a Norwegian financial services organization. Finally, in a recent study, Wagensveld (2013) focused on the travel and translation of the BSC in two Dutch organizations.

2.5 A comparison of how the three theoretical approaches explain the emergence and evolution of the BSC

The three theoretical views of the BSC have different areas of focus, but can offer complementary insights into how the BSC evolves in different local contexts. MFT-inspired BSC studies tend to focus primarily on macro-level dissemination and diffusion (Ax & Bjørnenak, 2005; Braam et al., 2007), and while recognizing the interpretive space of the concept on the supply side and demand side, these studies do not usually focus on how the concept is used at the micro level in organizations. By contrast, ANT researchers take a more micro-oriented view which emphasizes how important BSC actors enroll allies and build networks of support around the concept (Cooper et al., 2012; Qu, 2004). In addition, by focusing on the boundary-object characteristics of the BSC, ANT-inspired researchers highlight the plasticity of the BSC concept and how the concept is mobilized in organizational praxis in response to local organizational problems (Hansen & Mouritsen, 2005).

Finally, SIT-inspired researchers focus on how the BSC concept is translated and attains a local flavor as it is implemented in local organizations (Nilsen, 2007; Wagensveld, 2013; Wongkaew, 2007). The in-depth micro studies carried out by these researchers illustrate how the 'global' BSC concept is contextualized and attains a local flavor as it is implemented by users.

3. Research approach

The research reported in this paper was conducted as part of a larger project focusing on the impact of the BSC in Scandinavia (Madsen, 2011). The research approach was primarily qualitative and interpretive in nature. Interviews were the main data source in the study. A total of 18 actors involved in the Danish BSC market were interviewed, five of which were from the supply side (see Table 1) while 13 were users of the concept on the demand side (see Table 2). Interviewees were sought on the basis that they had experience of the BSC and an overview of how the BSC had evolved in Denmark. Conference brochures, media articles and websites were used to identify potential candidates. In some cases, a 'snowball procedure' was employed, where one interviewee recommended the next.

Table 1: Consultant interviewees

Interviewee	Firm size	Local/International	Firm specialization	Date
Consultant #1	Large	International	Generalist	26/01/2005
Consultant #2	Large	International	Generalist	27/01/2005
Consultant #3	Large	International	Generalist	08/04/2005
Consultant #4	Large	Local	Generalist	14/06/2005
Consultant #5	Small	Local	Specialist	15/06/2005

Table 2: User interviewees

Interviewee	Industry	Size	Local/international	Date
User #1	Consumer goods	Large	International	24/01/2005
User #2	Services	Large	International	28/01/2005
User #3	Financial sector	Large	Local	28/01/2005
User #4	Technology	Large	Local	31/01/2005
User #5	Accounting	Large	International	03/02/2005
User #6	Pharmaceuticals	Large	International	07/02/2005
User #7	Manufacturing	Large	International	17/03/2005
User #8	Manufacturing	Large	International	21/03/2005
User #9	Manufacturing	Large	International	11/04/2005
User #10	Financial sector	Large	Local	14/04/2005
User #11	Energy	Large	International	21/04/2005
User #12	Financial sector	Large	International	04/05/2005
User #13	Technology	Large	Local	30/05/2005

The interviews were semi-structured, lasted between 30 and 90 minutes, and covered several main themes – such as 1) how the actors came into contact with the concept, 2) their interpretations and approach, and 3) and their general experi-

ences from working with and implementing the concept. The interviews with the supply-side actors also focused on their views about the BSC market in Denmark, e.g. who the important actors have been, who they perceive as their main competitors, and how the local BSC market has developed over time. The interviews were recorded on an iPod, transcribed verbatim, and analyzed using a topic-centered approach comparing what different interviewees said about the respective topics.

4. Interpretations and use of the BSC: The supply-side perspective

This section reports on the interviews with suppliers of the BSC in Denmark. The interviews show that there is some variation concerning what the BSC means on the supply side, not only in terms of interpretations of the concept, but also regarding how the consultants draw on the concept in their own consulting work.

Table 3: How Danish suppliers interpret and mobilize the BSC concept

	Interpretation	Approach
C#1	Performance measurement system	Pragmatic approach; mostly using thoughts as a source of inspiration
C#2	Strategic management system	Close to KN's ideas
C#3	Strategic management system	Pragmatic approach; typically draws on the BSC as part of a 'larger package'
C#4	Strategic management system	Close to KN's ideas
C#5	Strategic management system	Close to KN's ideas

4.1. Interpretations

When asked how they interpret the BSC, four of five consultants said this was as a strategic management system, while one interpreted it as a performance measurement system. At the same time, the consultants underscored how flexible and loosely defined the BSC concept is. For example, one interviewee pointed out that:

> *»Balanced Scorecard is a very, very widely defined concept. You can fit a great deal into it.«*

The consultants argued that this characteristic can be useful when drawing on the concept in client projects. The BSC's interpretive space means that the concept can be used in many different ways and adapted to fit organization-specific circumstances. As one interviewee put it:

> *»The concept is so easy to understand and easy to adapt. You have some ideas that you can use in different ways.«*

4.2. Approach

With regard to how consultants approach the concept in practice, most point out that they try to stay close to KN's ideas and the BSC literature. However, at the same time, most of the consultants explained that they have adopted a rather pragmatic stance concerning the concept. Consider this quote from one consultant:

»I think I would use the thoughts behind the Balanced Scorecard as a source of inspiration.«

The BSC is also used pragmatically in the sense that few consultants were found to focus on the BSC as a 'standalone concept':

»We would not see this as a standalone component at all. I doubt that we would actually want to sell it. I think it needs to be followed by some other component before it will deliver any benefits.«

Instead, the BSC is often combined and mixed with other concepts in consultants' repertoires. Several consultants pointed out that they have incorporated BSC ideas into their consulting 'toolboxes' which usually comprise a range of different concepts and ideas. Yet, as one consultant noted:

»As an independent concept, I am not sure how many will continue to market it. I think there are many consultancies who use it as one element in their toolbox when working on project, but simply saying that you are going to do a Balanced Scorecard project, only Kaplan and Norton's own organization will do that. The rest of us have a broader approach to the concept.«

5. Interpretations and use of the BSC: The demand-side perspective

This section reports on the interviews with users of the BSC in Denmark. It includes illustrations of how the concept has been interpreted and used in Danish organizations.

Table 4: How Danish users interpret and mobilize the BSC concept

	Interpretation	Approach
U#1	Strategic management system	Close to KN's ideas
U#2	Performance measurement system	Close to KN's ideas; but used alongside the 'Service Profit Chain'
U#3	Performance measurement system	Pragmatic approach; tailor-made
U#4	Performance measurement system	Close to KN's ideas
U#5	Performance measurement system	Close to KN's ideas
U#6	Strategic management system	Close to KN's ideas
U#7	Performance measurement system	Pragmatic approach; BSC combined with the 'Business Excellence' concept
U#8	Performance measurement system	Pragmatic approach; BSC tailor-made
U#9	Performance measurement system	Pragmatic approach; BSC tailor-made
U#10	Performance measurement system	Pragmatic approach; BSC tailor-made
U#11	Performance measurement system	Close to KN's ideas
U#12	Performance measurement system	Close to KN's ideas
U#13	Performance measurement system	n/a

5.1. Interpretation

As with the consultants, several of the interviewees on the demand side acknowledge that the BSC concept has much room for interpretation, although the users did not appear to have reflected as much on this characteristic as the suppliers had. Overall, the organizations were positive in terms of the concept's properties and characteristics. For example, one manager pointed out that the BSC concept is an appealing one, saying: *»We liked the simplicity of it...«*

The BSC's scope for individual interpretation is also reflected in the day-to-day execution of the users studied. Most of the organizations interpret the BSC as a more comprehensive performance measurement system, close to the model originally put forth by KN in the 1990s. For example, one manager explicitly called his version of the BSC a 'measurement tool':

> *»Together with my assistant I developed a measurement tool that most people would call a Balanced Scorecard. We implemented this at the beginning of last year.«*

Similarly, another interviewee pointed out that although his organization is aware of more recent developments in the BSC literature, it has chosen to interpret the concept in the 'old-fashioned' way:

> *»If you look at Balanced Scorecard theory today where strategy maps and a broader view have been introduced, perhaps we represent more of a back-to-basics approach...Our methods around the Balanced Scorecard are closer to the original, introduced in the early 1990s, in that we have not adopted the new ideas.«*

Only two of the user organizations interpreted the BSC primarily as a strategic management system. KN's more recent ideas such as mapping of cause-and-effect relationships and strategy maps were seldom mentioned by the interviewees. Some managers did not appear to be aware of these aspects of BSC theory, while other organizations admitted that although they were aware of reference to a strategy map and so on, they had not placed much emphasis on these aspects due to limited time and resources. Consider the following comment:

> *»When it comes to cause-and-effect relationships we have a missing link. We haven't spent much time on that. We ought to have done.«*

5.2. Approach

Many of the interviewees reported that they have tried to stay close to the BSC 'theory' developed by KN. They tend to use the standard four perspectives or some variation of them. Often, they will change some terms for a better fit with existing practices and language used in their respective organizations. As one user representative put it,

»We use more or less the original model, but we use other terms. When we started in 1998-1999 some terms were already being used in the organization, and we didn't want to change these just because we had drawn inspiration from Kaplan and Norton. Instead we wanted to stick to our traditional language.«

Other organizations identify themselves explicitly as 'pragmatic' users of the BSC. In our study, these candidates indicated that they have a very pragmatic stance towards the concept. These organizations have moved away from the 'complete BSC' and are only using certain aspects of the concept (cf. Nielsen & Sørensen, 2004). For instance, two organizations indicated that initially they had been very faithful to the concept as described by KN, but that gradually they had become more pragmatic:

»Some organizations are using the Balanced Scorecard in its original form, but we have made a tailor-made version. We are not faithful to the theory, we really couldn't care less... Our concept is very tilted. In the beginning it was 100% theory, but after a year or so we discovered that it was 'junk' in many ways because we did not have an organization that could utilize it. So we have chosen to tilt it towards areas where we could reap the benefits... At the start we had a standard specification, i.e.: 'This is the Balanced Scorecard.' After a while we became aware that 'We are not Balanced, and so what!'

»From having a 100% theory-faithful and orthodox implementation, we have moved toward a more organization-specific model based on the Balanced Scorecard. We now have a more liberal attitude towards the theory. It is a part of the process, going from faithfulness to being more independent. But the premise behind the Balanced Scorecard has a future in our organization.«

In a similar vein, one organization explained how over time it had moved away from using what it referred to as a 'textbook BSC' towards its own adaptation of the BSC which focuses more on traditional financial indicators:

»I believe that our use of the BSC is a bit different from the original model and we are turning to a concrete focus on financial parts of the BSC. We call our approach a 'pick and choose'.«

6. Discussion

This section will discuss the findings in relation to the three theoretical perspectives on the diffusion, translation and popularization of the BSC presented in the first part of this paper. First, we will discuss the characteristics of the BSC and their implication. Second, we will examine how the BSC has been shaped in a Danish context. Finally, we will look at the mismatch between how the concept is mobilized on the supply versus the demand side.

6.1. The characteristics of the BSC and their implications

In the theoretical part of this work, the interpretive space and boundary objects characteristics of the BSC were identified as a particularly important in explaining how the BSC is interpreted and used in practice. The interpretive space is also reflected in Danish practice where the concept is interpreted and used in different ways. Different versions of the BSC exist in practice as the concept is mobilized in relation to organization-specific problems and challenges (cf. Hansen & Mouritsen, 2005). However, although the concept is 'plastic', certain elements of the concept remain relatively stable across the different organizations. For example, most organizations explain that they apply some variation of the standard four perspectives and a combination of lead-and-lag indicators. Hence, the BSC can be considered a boundary object since it retains its 'core' as it is diffused, translated and popularized in different settings (Cooper et al., 2012; Hansen & Mouritsen, 2005; Wagensveld, 2013).

6.2. How the BSC concept has been shaped in the Danish context

The data shows that the BSC concept has been shaped by both suppliers and users in a Danish context. In the theory section of this paper we have identified several typologies of BSC use which distinguish between interpretations of the BSC, ranging from use as a measurement system to a strategic management system (Brudan, 2005; Lawrie & Cobbold, 2004; Speckbacher et al., 2003). The interview data show that there are two main interpretations of the BSC in practice, i.e. its use as a 'performance measurement system' or as a 'strategic management system'. However, the data also illustrate how these interpretations are mobilized in various ways in practice on both the supply and the demand side, as actors draw on the concept pragmatically. The concept is combined and mixed with other concepts and ideas, resulting in a multiplicity in terms of how the concept is interpreted and used.

6.3. 'Mismatch' between how the concept is mobilized on the supply- and the demand-sides

The data show a mismatch between the way suppliers and users interpret and attach meaning to the concept of the BSC. *Suppliers* of the concept tend to view and portray the concept as a 'strategic management system', highlighting more recent developments in the BSC literature such as causal relationships and strategy maps. However, *users* of the concept tend to interpret it mainly as a 'performance measurement system'. This last finding can be seen in the light of other studies which have noted that the many users of the BSC engage the concept primarily as a performance measurement system (Ax & Bjørnenak, 2005; Speckbacher et al., 2003).

7. Conclusion

The purpose of this paper has been to examine how Danish suppliers and users have interpreted and attached meaning to the concept of the Balanced Scorecard. The data show that there is a multiplicity of ways in which the BSC is interpreted and mobilized in praxis. The interviews we have undertaken suggest that actors use the concept's interpretive space and boundary-object characteristics to mobilize and attach meaning to the BSC. The actors draw on the BSC rather

pragmatically and combine and blend it with other concepts and ideas, resulting in organization-specific manifestations of the BSC. One interesting finding is the mismatch in how suppliers and users interpret the concept. While suppliers tend to promote the BSC as a 'strategic management system', users tend to view it as a 'performance measurement system'.

7.1. Contributions

It is maintained that the current paper contributes to the literature on the diffusion and popularization of the BSC in two ways. First, it shows variation in the way different types of actors interpret and attach meaning to the BSC. The combination of supply- and demand-side data provides a more 'balanced' view of how the concept has impacted praxis. In addition, the combination of supply- and demand-side data has allowed for the identification of a mismatch in how suppliers and users of the concept interpret and attach meaning to the BSC, which would not have been possible if data had been gathered from only one side of the market.

Second, the paper blends insights from three different theoretical perspectives (MFT, ANT, and SIT) on how the BSC concept is diffused, translated and popularized. These three views can illuminate different aspects of how the BSC evolves in local contexts, but one point of convergence is the 'interpretive space' and 'boundary-object' characteristics of management concepts.

7.2. Limitations

The paper has several limitations and shortcomings which should be noted. First, it is possible to argue that the data utilized in this study was not sufficiently 'micro', given that only one representative from each organization was interviewed. In future studies researchers should consider interviewing a range of different representatives from organizations and build more in-depth case studies, e.g. similar to the study by Hansen and Mouritsen (2005). However, there is a trade-off between depth and breadth, particularly if one wants to say something meaningful about the concept's reception not only at the micro level, but also at the macro (i.e. national) level.

Second, the theoretical eclecticism in the paper may attract criticism from theoretical purists who would argue that MFT, ANT, and SIT are ontologically distinct. Recently, however, researchers have started to combine different approaches such as neo-institutional theory and ANT (Lounsbury, 2008). Hence, it can be argued that theoretical eclecticism is important in order to grasp the complexity of processes related to diffusion, translation and popularization.

Third, the mismatch between how the concept is mobilized on the supply side and the demand side could be an 'early-phase phenomenon'. The data were gathered in the mid-2000s when the concept of the BSC was still relatively new, and more recent terms such as 'strategy maps' were still not much used in organizational praxis. Hence, it is not inconceivable that if the study were conducted today the

level of knowledge about the BSC would more evenly distributed among suppliers and consumers, which would influence interpretations and mobilizations of the concept. Also, it is likely that later adopters (e.g. post-2005) would be more likely to be aware of and implement the tools and methodologies described in more recent BSC books.

7.3. Future work

The paper's limitations should be addressed in future work. For example, in-depth case studies could provide a deeper understanding of how the BSC is interpreted and mobilized in relation to organization-specific problems and issues (cf. Hansen & Mouritsen, 2005). Such studies would require more micro data, e.g. by interviewing multiple informants at different levels of the organization.

In order to provide a more complete overview of how the concept of the BSC has been received in a Danish context, it would be an advantage to combine not only qualitative data about organization-specific interpretations and translations, but also quantitative data (e.g. survey data) which might give a broader overview of how the concept is used in a broader population of organizations in Denmark (cf. Benders & Van Bijsterveld, 2000).

Finally, future studies could also provide a more updated picture of how the concept is currently interpreted in a Danish setting. Such studies could provide insight into how the concept evolves as a practice within a Danish context over time, and the various contextual factors shaping interpretations and use of the concept.

References

Abrahamson, E. 1996. Management Fashions. *Academy of Management Review*, 21: 254-285.

Aidemark, L.G. 2001. The Meaning of Balanced Scorecards in the Health Care Organization. *Financial Accountability & Management*, 17(1): 23-40.

Ax, C., & Bjørnenak, T. 2005. Bundling and diffusion of management accounting innovations – the case of the balanced scorecard in Sweden. *Management Accounting Research*, 16: 1-20.

Benders, J., & Van Bijsterveld, M. 2000. Leaning on lean: The reception of management fashion in Germany. *New Technology, Work and Employment*, 15: 50-64.

Benders, J., & Van Veen, K. 2001. What's in a fashion? Interpretive viability and management fashions. *Organization*, 8(1): 33-53.

Benders, J., & Verlaar, S. 2003. Lifting parts: putting conceptual insights into practice. *International Journal of Operations & Production Management*, 23(7): 757-774.

Braam, G. 2012. Balanced Scorecard's Interpretative Variability and Organizational Change. In C.-H. Quah, & O.L. Dar (Eds.), *Business Dynamics in the 21st Century*: InTech.

Braam, G., Benders, J., & Heusinkveld, S. 2007. The balanced scorecard in the Netherlands: An analysis of its evolution using print-media indicators. *Journal of Organizational Change Management*, 20(6): 866-879.

Braam, G., Heusinkveld, S., Benders, J., & Aubel, A. 2002. The reception pattern of the balanced scorecard: Accounting for interpretative viability, *SOM-Theme G: Cross-contextual comparison of institutions and organisations*. Nijmegen, The Netherlands: Nijmegen School of Management, University of Nijmegen.

Braam, G., & Nijssen, E. 2004. Performance effects of using the Balanced Scorecard: a note on the Dutch experience. *Long Range Planning*, 37: 335-349.

Briers, M., & Chua, W. F. 2001. The role of actor-networks and boundary objects in management accounting change: a field study of an implementation of activity-based costing. *Accounting, Organizations and Society*, 26(3): 237-269.

Brudan, A. 2005. Balanced Scorecard typology and organisational impact. *KM Online Journal of Knowledge Management*, 2(1).

Carson, P., Lanier, P., Carson, K., & Guidry, B. 2000. Clearing a path through the management fashion jungle. *Academy of Management Journal*, 43(6): 1143-1158.

Christensen, K.S., & Bukh, P.N. 2013. *Succes med balanced scorecard*. Copenhagen, Denmark: Gyldendal A/S.

Clark, T. 2004. The fashion of management fashion: A surge too far? *Organization*, 11(2): 297-306.

Cooper, D.J., Ezzamel, M., & Qu, S. Q. 2012. Popularizing a Management Accounting Idea: The Case of the Balanced Scorecard, *AAA 2012 Management Accounting Section (MAS) Meeting Paper*.

Czarniawska, B., & Sevón, G. 2005. *Global ideas: How ideas, objects and practices travel in the global economy*. Danmark: Copenhagen Business School Press.

Czarniawska, B., & Sevòn, G. 1996. *Translating organizational change*. Berlin: Walter de Gruyter.

Dechow, N. 2012. The Balanced Scorecard-Subjects, Concept and Objects-A Commentary. *Journal of Accounting & Organizational Change*, 8(4): 511-527.

Giroux, H. 2006. 'It was such a handy term': Management fashions and pragmatic ambiguity. *Journal of Management Studies*, 43(6): 1227-1260.

Hansen, A., & Mouritsen, J. 2005. Strategies and Organizational Problems: Constructing Corporate Value and Coherences in Balanced Scorecard Processes. In C. S. Chapman (Ed.), *Controlling Strategy: Management, Accounting and Performance Measurement*: 125-150. New York: Oxford University Press.

Hoque, Z. 2014. 20 years of studies on the Balanced Scorecard: Trends, accomplishments, gaps and opportunities for future research. *The British Accounting Review*, 46(1): 33-59.

Jung, N., & Kieser, A. 2012. Consultants in the Management Fashion Arena. In M. Kipping, & T. Clark (Eds.), *The Oxford Handbook of Management Consulting*: 327-346. New York, USA: Oxford University Press.

Justesen, L., & Mouritsen, J. 2011. Effects of actor-network theory in accounting research. *Accounting, Auditing & Accountability Journal*, 24(2): 161-193.

Kaplan, R.S., & Norton, D.P. 1992. The balanced scorecard – Measures that drive performance. *Harvard Business Review*(January-February): 71-79.

Kieser, A. 2002. Managers as Marionettes? Using Fashion Theories to Explain the Success of Consultancies. In M. Kipping, & L. Engwall (Eds.), *Management Consulting – Emergence and Dynamics of a Knowledge Industry*: Oxford.

Latour, B. 1987. *Science in action: How to follow scientists and engineers through society*: Harvard university press.

Lawrie, G., & Cobbold, I. 2004. Third-generation balanced scorecard: evolution of an effective strategic control tool. *International Journal of Productivity and Performance Management*, 53(7): 611-623.

Lounsbury, M. 2008. Institutional rationality and practice variation: New directions in the institutional analysis of practice. *Accounting, Organizations and Society*, 33(4): 349-361.

Madsen, D.Ø. 2011. *The impact of the balanced scorecard in the Scandinavian countries: a comparative study of three national management fashion markets*. Unpublished Ph.D. Thesis, Norwegian School of Economics, Department of Strategy and Management, Bergen, Norway.

Malmi, T. 2001. Balanced scorecards in Finnish companies: A research note. *Management Accounting Research*, 12: 207-220.

Marr, B. 2005. Performance measurement: an overview of the current state of use in the USA. *Measuring Business Excellence*, 9(3): 56-62.

Modell, S. 2004. Performance Measurement Myths in the Public Sector: A Research Note. *Financial Accountability & Management*, 20(1): 39-55.

Modell, S. 2009. Bundling management control innovations: A field study of organisational experimenting with total quality management and the balanced scorecard. *Accounting, Auditing & Accountability Journal* 22: 59-90.

Modell, S. 2012. The Politics of the Balanced Scorecard. *Journal of Accounting & Organizational Change*, 8(4): 475-489.

Nielsen, S., & Sørensen, R. 2004. Motives, diffusion and utilisation of the balanced scorecard in Denmark. *International Journal of Accounting, Auditing and Performance Evaluation*, 1: 103-124.

Nilsen, E.A. 2007. *Oversettelsens mikroprosesser. Om å forstå møtet mellom en global idé og lokal praksis som dekontekstualisering, kontekstualisering og nettverksbygging*. University of Tromsø, Tromsø.

Niven, P.R. 2005. *Balanced scorecard diagnostics: Maintaining maximum performance*. Hoboken, NJ: : John Wiley & Sons, Inc.

Nørreklit, H. 2003. The balanced scorecard: What is the score? A rhetorical analysis of the balanced scorecard. *Accounting, Organizations and Society*, 28(6): 591-619.

Powell, W.W., Gammal, D. L., & Simard, C. 2005. Close Encounters: The Circulation and Reception of Managerial Practices in the San Francisco Bay Area Nonprofit Community. In B. Czarniawska, & G. Sevon (Eds.), *Global Ideas*: CBS Press.

Qu, S. 2004. A sociological analysis of the rise and dissemination of management accounting innovations: Evidence from the balanced scorecard, *Fourth Asia Pacific Interdisciplinary Research in Accounting Conference*. Singapore.

Rigby, D., & Bilodeau, B. 2013. Management Tools & Trends 2013. London: Bain & Company.

Røvik, K.A. 1998. *Moderne organisasjoner*. Oslo: Fagbokforlaget.

Røvik, K.A. 2002. The secrets of the winners: Management ideas that flow. In K. Sahlin-Andersson, & L. Engwall (Eds.), *The Expansion of Management Knowledge: Carriers, Ideas and Sources*, Vol. 113-144. Stanford: Stanford University Press.

Sahlin-Andersson, K., & Engwall, L. 2002. *The expansion of management knowledge: Carriers, flows and sources*. Palo Alto, California, USA: Stanford University Press.

Sahlin, K., & Wedlin, L. 2008. Circulating ideas: Imitation, translation and editing. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin-Andersson (Eds.), *The SAGE Handbook of Organizational Institutionalism*. Thousand Oaks: Sage.

Sibbet, D. 1997. 75 years of management ideas and practice 1922-1997. *Harvard Business Review*, 75(5): 2-12.

Silk, S. 1998. Automating the balanced scorecard. *Management Accounting*(May): 38-40, 42-44.

Speckbacher, G., Bischof, J., & Pfeiffer, T. 2003. A descriptive analysis of the implementation of balanced scorecards in German-speaking countries. *Management Accounting Research*(December): 361-388.

Star, S.L., & Griesemer, J.R. 1989. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3): 387-420.

Strang, D., & Meyer, J.W. 1993. Institutional conditions for diffusion. *Theory and Society*, 22(4): 487-511.

ten Bos, R. 2000. *Fashion and utopia in management thinking.* Amsterdam/Philadelphia, PA: John Benjamins.

Wagensveld, J. 2013. *The travel and translation of balanced scorecards*. Radboud Universiteit Nijmegen.

Wongkaew, W. 2007. *Managing Multiple Dimensions of Performance: A Field Study of Balanced Scorecard Translation in the Thai Financial Services Organisation*. University of Warwick, UK.


This paper focuses on how
Modeling (SDM) can
ased understanding of
tics' setup. Such
necessary to identify
effectively support a
Balanced Scorecard (BSC).
idea of a BSC as a
systems thinking, the
address not only
cept of 'comprehensiveness'
i.e. the 'scope' (learning)
(system dynamics causality
also the methodological
'precision' (solution by
We show that a company
derstanding of its
stage in the BSC design
such as causal loop
ceive management
vironmental

Management Accounting and Business Analytics

An example of System Dynamics Modelling's use in the design of a Balanced Scorecard

Steen Nielsen, Erland H. Nielsen, Anders Jacobsen and Lars Bjørn Pedersen[1]

Abstract

This paper focuses on how System Dynamics Modeling (SDM) can contribute to an increased understanding of business in an 'analytics' setup. Such improved understanding is necessary to identify the variables that most effectively support a company's strategy in a Balanced Scorecard (BSC). By combining the idea of a BSC as a causal loop system with systems thinking, the current paper aims to address not only the conceptual domain concept of 'comprehensiveness' related to BSC, i.e. the 'scope' (learning) and 'differentiation' (system dynamics causality and feedback), but also the methodological domain concept of 'precision' (solution by differential equations). We show that a company can improve its understanding of its business at a very early stage in the BSC design phase by using SDM such as causal loop diagrams.

1. Introduction

Within management accounting a frustration point has been that few research results have ever been used in the practical world (Merchant, 2012), even though management accounting is actually an applied and practical field that continually faces new challenges from the business world in real life (Kasanen *et al.,* 1993; Kaplan, 1998, 2012; Otley, 2001; IMA, 2008; CIMA, 2009; Merchant, 2012). Having research contribute to new theories does not suffice; rather researchers should try to advance the body of knowledge. This means developing theories and approaches that specify both the behavior and the context required for achieving specified outcomes (Kaplan, 2011; Merchant, 2012; Ahrens, and Chapman, 2007). In other research areas, there has been intense discussion about the concept of 'analytics' and its practical implications for areas such as marketing, logistics and human resources (Liberatore and Luo, 2010; Davenport, 2006).

Analytics can be defined as the use of business analytics, business intelligence, data analytics, big-data analytics and analytical skills (Davenport and Harris, 2007). Several surveys have shown the importance of analytics for the future (e.g. Accenture, 2013; McKinsey, 2011); surveys within management accounting too have shown that it will become increasingly important for management ac-

countants to have analytical skills in future (ACCA, 2009; CIMA, 2008) as well as knowledge of quantitative tools (CIMA, 2009; Collier *et al.*, 2007). Even though the concept of 'big data' or 'big-data analytics' is often applied in a marketing context, it is also applicable to performance management (Eckerson, 2011, Giles, 2012 in The Economist).

In a 2008 interview with Paul Sharman, the President of the American IMA (Institute of Management Accountants), Kaplan pointed out too that *'Management accounting analytics are no longer constrained by limited or complex access to companies' databases. But to excel at analytics, management accountants will require extensive training in modeling, multivariate statistics, and econometrics.'* The tendency is also seen in the field of New Public Management, i.e. the use of accounting techniques within the public sector (see e.g. IBM, 2011; SAS Institute White paper, 2011).

An exhaustive search of published literature on management accounting/control and (business) analytics by the authors, across a number of library databases (e.g. ABI/INFORM Global, Business Source Complete) and on the internet (e.g. Google, Business Source Complete and Social Science Research Network), yielded only a handful of papers discussing the convergence of business analytics and management accounting (e.g. in Nielsen *et al.,* 2010; Silvi *et al.,* 2013).

Over the last few years, surveys have shown increased interest in using the BSC as a holistic planning tool (Otley, 2003; De Geuser *et al.,* 2009; Wiersma, 2009) with a feed-forward as well as a feedback model, often called the 'Closed-Loop Management System' (Kaplan & Norton, 2008).

Inspired by these tendencies we formulated the following research question: '*How can a company use a system dynamics approach to accomplish the development of its analytical PMS (Performance Management System) in order to increase the insight of organizational learning?*' The structure of this research process is inspired by Brinberg and McGrath's (1985) Validity Network Schema (VNS) in which research contains relations between domains, levels, stages and paths.

As Sterman eloquently puts it, *'Change is accelerating, and as the complexity of the systems in which we live grows, so do the unanticipated side-effects of human actions, further increasing complexity. Many scholars call for the development of systems thinking to improve our ability to manage wisely. But how do people learn in and about complex dynamic systems? Learning is a feedback process in which our decisions alter the real world. We receive information feedback about the world and revise the decisions we make and the mental models that motivate those decisions'* (J. D. Sterman, 1994, Article first published online: 26 December 2006).

In stage one, we seek to combine data and information from the company (substantive domain) with the methodological domain (the clausal loop diagrams

techniques from system dynamics) in order to identify a number of causal mechanisms and relations.

'Effective methods for learning in and about complex dynamic systems must include (1) tools to elicit participant knowledge, articulate and reframe perceptions and create maps of the feedback structure of a problem from those perceptions; (2) simulation tools and management flight simulators to assess the dynamics of those maps and test new policies; and (3) methods to improve scientific reasoning skills, strengthen group process and overcome defensive routines for individuals and teams' (J. D. Sterman, 1994, see reference above).

The outcome is a specific theory of the company based on system dynamics definitions and assumptions. Sections 2 and 3 describe this stage. The next step is to form a number of differential equations in the Vensim™ software to enable us to quantify and simulate a number of scenarios in system dynamics. The outcome is shown in section 4.

2. Integration of the Balanced Scorecard with system dynamics

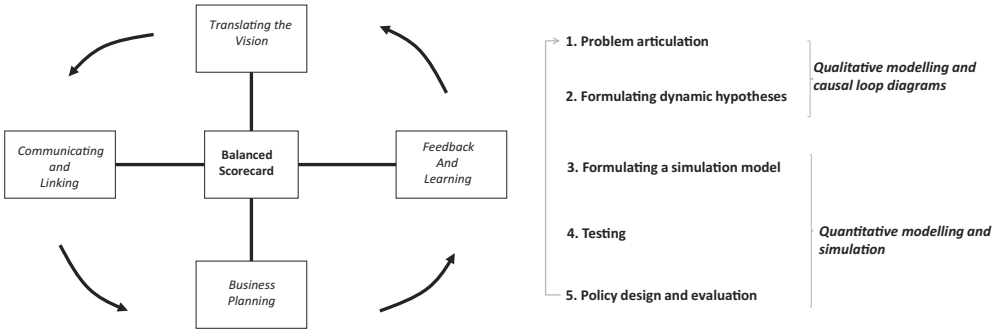
Since 1992 (Kaplan and Norton,1992) the Balanced Scorecard (BSC) has evolved through different developments and versions. A huge number of survey and case studies have been conducted (e.g. Malina and Selto, 2001). In Denmark, the BSC model is also a popular framework for performance management (Nielsen and Sørensen, 2004; Sandalgaard and Bukh, 2008). The idea of such a 'holistic framework' for a management control system is not new, however (see e.g. Anthony, 1965; Flamholtz, 1979; Lyneis, 1980).

Despite its popularity and appealing simplicity at a conceptual level, the BSC concept has also been the subject of criticism. Criticism and discussion of the BSC can be referred to as a 'methodology disagreement' or – more specifically – a disagreement about how to estimate time-lags (e.g. between different perspectives) and the causes and effects of KPIs (key performance indictors). Several attempts have been made to see whether (traditional) statistical significance might be established (Malina *et al.,* 2007) based on Sims-Granger type causality (Granger, 1969). However, others have argued that the classical statistical view and assumptions on cause and effect are mistaken or not applicable here, and that alternative interpretations should be used (Bukh and Malmi, 2005; Wallenburg and Weber, 2006; Kaplan, 2012). Or, as Meadows (1980) puts it, *'The non-linear feedback structure of system dynamics models makes standard econometric techniques either inapplicable or extremely difficult to use'* (for an example see Sterman *et al.,* 1997 in combination with TQM). This may be caused by the fact that deep, 'nature-given' causal relations are almost impossible to encounter in socio-economic systems (Senge and Forrester, 1980). Instead, the concept of mental models as discussed within system dynamics (Forrester, 1958; Warren, 2004; Senge, 1990) seems much more appropriate for BSC modeling in quantitative terms.

System dynamics thinking was first introduced by Forrester in 1958 as an example of a macroeconomic problem of production distribution. Applications have now expanded to environmental change, politics, economic behavior, medicine and other fields (Lane, 1992; Vennix, 1999; Warren, 2008). The main idea is that 'you can't just do one thing' without thinking that 'everything is connected to everything else', which is also at the center of analytical performance management.

Inspired by Kaplan and Norton (2007) and Sterman (2000), Figure 1 shows steps and points of resemblances for SDM (System Dynamic Modeling) and a BSC.

Figure 1: The Balanced Scorecard stage model and steps in system dynamics



A number of studies have been conducted using system dynamics for performance management/measurement based on different assumptions (Akkermans and Oorschot, 2005; Nielsen and Nielsen, 2008; Capelo and Dias, 2009; Salterio, 2012).

3. Case study

The subject for our case study is a service department within an international company. In the autumn of 2009 the company began to design and work with the Balanced Scorecard concept with a view to optimizing strategy execution throughout the organization. Consultants from the Palladium Group were used as sparring partners in the implementation and the Balanced Scorecard Group within the company was given the primary responsibility for implementation.

The formulation of the strategy was the first step and had already been conducted for the department. When we started our study, the department worked with the setup of a strategy map and strategic objectives (or what Spreckbacher *et al.* (2003) call a type I BSC).

The first two steps in Figure 1 are the qualitative issues, i.e. the formulation of the problem in formal terms and the formulation of a dynamic hypothesis (Sterman, 2000).

3.1. Formulation of the modeling problem

The formulation of the problem is seen as the most important step (Forrester, 1991; Sterman, 2000; Warren, 2008). Here, reference modes, selection of key variables, the time over which the model should be used, and the settings for the dynamic problem formulation are set out. The primary tasks for the team were the formulation of the problem, mapping the interaction between the KPIs in the support department in order to bring its efficiency to the top-of-world class as seen from the department's view and as expressed in the strategy.

The department had already defined and established a large number of key variables and various efficiency measures, some of which were filtered out as part of the iterative modeling process in order to achieve a more focused initial situation. In partnership with the department staff we identified and selected three key variables as the main drivers supporting the purpose of this study. These were *'user recommendation of the department'*, which is a measure of whether users are satisfied with the service supplied; '*workload'*, i.e. a measurement for the number of cases (work assignments) that are not solved in the time allowed; and '*perceived competence fit'* which measures whether employees possess the necessary skills to solve the daily tasks set by other parts of the organization.

3.2 Formulation of the dynamic hypotheses and the causal loop diagrams

When the purpose of the model was identified and characterized over an appropriate time horizon a 'theory' was developed (also called the dynamic hypothesis) to account for the problematic behavior of the model.

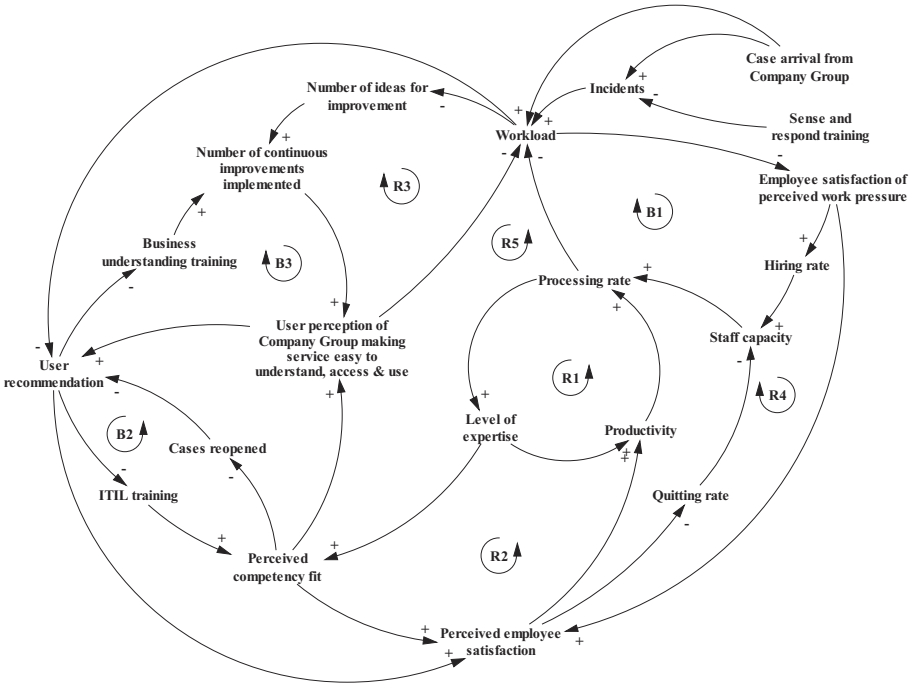
Finally, the Causal Loop Diagram (CLD) had to be developed. This is based on the model boundary chart (not shown here) and illustrates the interactions between the identified key variables. The CLD diagram is shown in Figure 2 and is the result of the discussion with the department.

The CLD is identified through eight interacting feedback loops (where + and – indicate the type of change-based causal reaction ('correlation') to be expected), which together exemplify the model's dynamic nature. The loops are either referred to as »R« or »B«, referring to self-reinforcing loops (Reinforcing) and balanced loops (Balancing)[2].

3.3 The quantitative structure of the System Dynamics Model

We used the qualitative model as a skeleton for a System Dynamics Model (SDM) for the BSC. We used the department team as sparring partners and gave them what Richardson and Andersen (1995) refer to as a gatekeeper role, providing valuable information and learning. A final finesse of the closed-loop causal reasoning described above is that it relates almost directly to the SDM tradition as it has been implemented for example in the VENSIM™ software we used for our project. Below, we show only a single CLS example from the process perspective (the incoming case subsystem) and value prospective (the efficiency subsystem) based on stocks and flows.

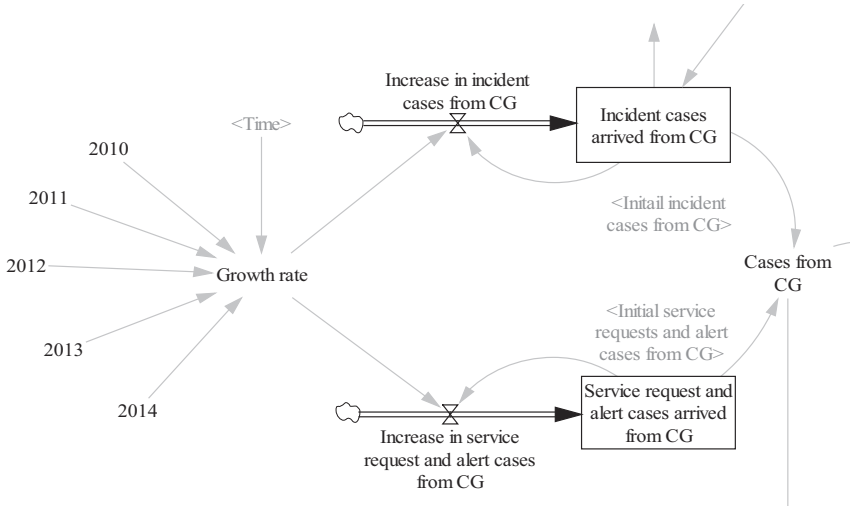
Figure 2: The causal loop diagram



3.3.1 Process perspective

The service department categorizes cases in three groups, with special focus given to incidents representing something out of the ordinary. Figure 3 shows the number of incoming cases and, as can be seen, unusual incidents are recorded separately whereas the other two types of cases are recorded together. Trends in the number of incoming cases are obtained through two stocks where the initial value is the number of cases recorded in 2009. These stocks are multiplied by an estimated growth rate for each of the five years of the simulation time. The annual growth rate used in the model is estimated by the department team and is an expression of the mental image rather than the result of estimated forecast. As the final step, the numbers of the three case types are added together to form a single variable.

Figure 3: Incoming case subsystem



It is important to realize that this perspective consists of both a capacity and a productivity module (not shown here). It is based on learning-curve theory (see e.g., Hirsch, 2002, chapter 5), in which experience is linked with productivity. We have used the model provided by Sterman (2000, p. 507) for our calculations of productivity.

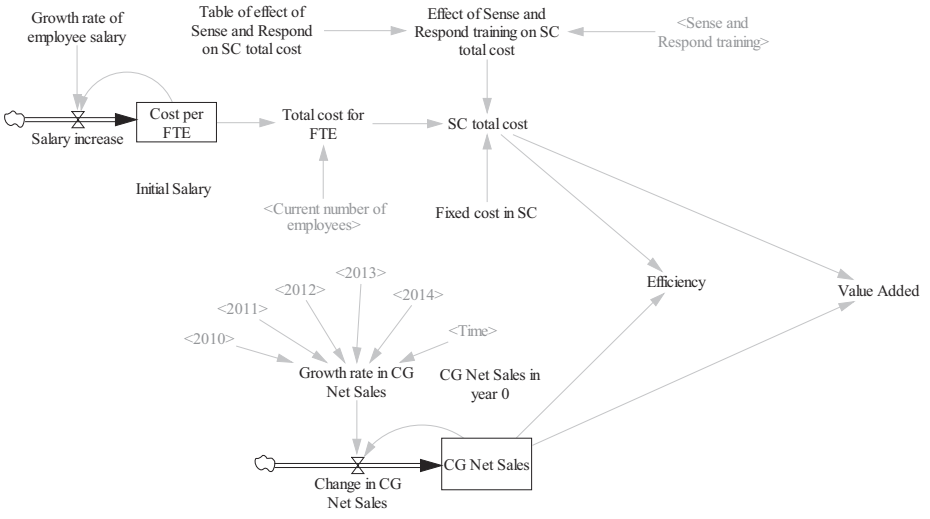
3.4. Value-add perspective

The department does not get a specific financial payment for the service it provides to other departments; it is instead measured on the change created by the department in the net sales for the company group as a whole. Because the two all-important goals are effectiveness and efficiency, the SDM is designed to focus on and support these goals. The efficiency target focuses on the department's ability to keep costs down, while the effectiveness target is geared to the level of service provided.

As the department has a supporting role, the actual costs of the department have to be compared to planned costs to give a true and fair view of the development over time. *Efficiency* is measured as the service department's total costs relative to the total net sales for the whole company, whereas *value-add* is measured as the company's total net sales minus the total costs of the department.

The efficiency subsystem is shown in Figure 4.

Figure 4: The efficiency subsystem



The fixed costs are set to be constant throughout the simulation period because changes occur primarily in variable costs. The variable costs are accounted for predominantly by labor expenses. The model includes an annual wage increase for employees of 3%. We have set a specific time-span for the growth rates (and the setting for the simulation) as being from the start of 2010 to the end of 2014.

4. Validation, experiments and results

Validation of System Dynamics Models has two major aspects: *structure* validation and *behavior* validation (Barlas 1996; Sterman, 2000; Warren, 2008). *Structure validation* aims to demonstrate that the model's internal structure (set of relationships) is an adequate description of the real system with respect to the problems of interest. *Behavior validation* aims to check whether the output behavior of the model adequately describes real dynamic behavior. The structural validation of the service department's model was carried out by a range of logic tests, extreme-condition, sensitivity and boundary tests, which will not be discussed here. We simply state that the model was found to be structurally reliable and also expressed behavioral validity.

Numerous simulation experiments in Vensim™ (Eberlein and Peterson, 1992) were carried out on the model in order to find answers to the problems formulated at the beginning of this paper, and to predict a few possible outcomes based on finding relevant KPIs that strongly affect the company's strategy execution (De Geus, 1992; Simon, 1990). For the purposes of this study we will only change two input variables within the Learning and Growth Perspective to see the changes in six KPIs from different perspectives as outcome variables. The scenarios are:

- Base: simulation of the current model with original parameters.
- Base1: simulation of an increase in '*multiplier for employees' satisfaction*' of 20 %.
- Base2: simulation of an increase in '*multiplier for perceived competence fit*' of 20 %.

In the short term, if the company wants to improve the result (meaning improvement in user recommendation and cost efficiency), the choice is either an increase in employee satisfaction (Base1) or an increase in employees' competence (Base2). Optimization by a mix of the scenarios might also be possible, but will not be done here. The chosen KPIs are all so-called »lagging« KPIs in the BSC setup, meaning that they accumulate information from other KPIs within the same perspective for passing information to new KPIs and perspectives (Kaplan and Norton, 1996).

Figure 5: KPI lag-indicators from a Learning and Growth Perspective

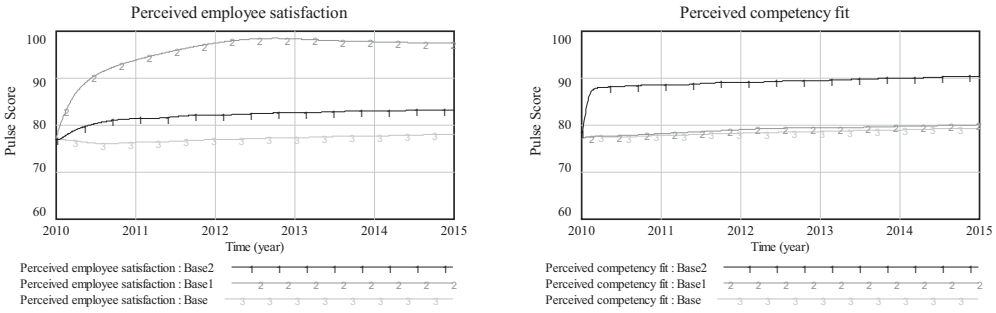


Figure 5 illustrates how the two KPIs increase over the timescale for the simulation. Note that *perceived employee satisfaction* does not rise immediately as employees' perception of satisfaction is not detected immediately. The model explicitly illustrates these delays, which is one of the great advantages of using SDM for BSC. Therefore, if the department launched an initiative on employee satisfaction (e.g. job rotation), System Dynamics Modeling would make it possible to understand the dynamic mechanisms and then the delayed impact on user satisfaction would not come unexpectedly. The model outcome clearly shows that the full effect of the increase in employee satisfaction in the department can only be read in user satisfaction after about 2.5 years, meaning that the financial result comes even later. Besides, the identification of delays in the KPIs in the model also contributes to an understanding of the side-effects, i.e. an increase in *perceived competence fit*. Depending on the construction of the model's dashboard it is possible to illustrate that an increase in employee satisfaction also positively affects several other KPIs.

Figure 6: KPI lag-indicators from a Process Perspective

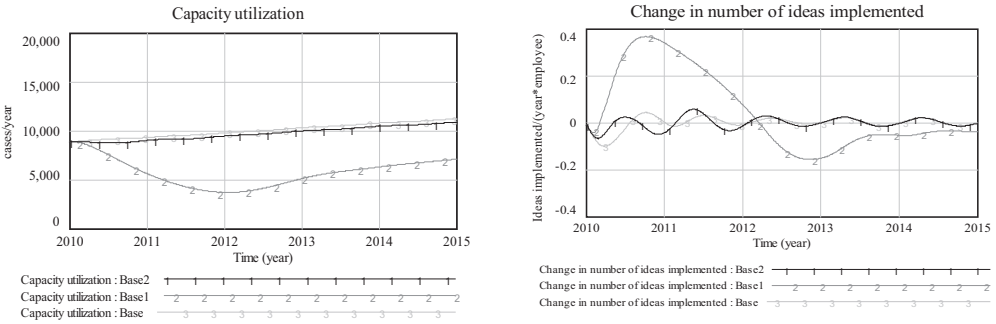
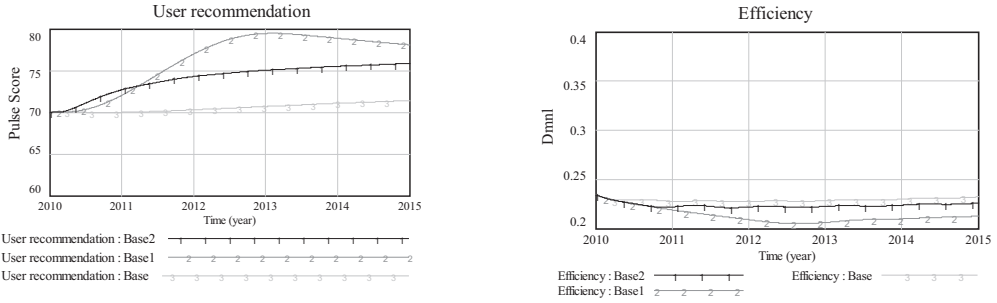


Figure 6 shows the simulation output for two KPIs from the process perspective. Both *capacity utilization* and *changes in number of ideas implemented* change markedly when the company is able to increase the satisfaction of the employees by 20% (Base1). For *capacity utilization* this implies that the company is able to come up with free capacity that may be used for other things (e.g. insourcing). At the same time innovation is initiated through an increase in the number of ideas implemented. Also note that »free capacity« and »number of new ideas implemented« (i.e. innovation) change at different points in time in the simulation, i.e. they are both delayed, but at different rates (see Figure 6). We also see that satisfaction is a much better idea than better competences (compare Base2 and Base1). As a result of the multiplier effect the KPI quickly rises to a high level.

Figure 7: KPI lag-indicators from a Value-Added Perspective



Two KPIs from the value-added perspective are shown in Figure 7. The graph for *user recommendation* shows that the effect of employee satisfaction (Base 1) comes rather late, almost three years after the initiative was brought in. The (cost) *efficiency* graph further shows that additional efficiency results from an increase in employees' satisfaction with 20% Base1) rather than from an increase in *competence fit* (Base2). *User recommendation* indicates the effectiveness of the department as this KPI measures the demand for its services. Note that the development of the two KPIs is offset as the delays are included in the model.

5. Conclusion

This paper used the Brinberg and McGrath (1985) research setup in combination with the idea of 'analytics' to document the possibilities of designing and using the SDM approach for the Balanced Scorecard model. To further focus our model we also used two other themes that are often discussed in relation to system dynamics: the Double-Loop Learning (Senge, 1990; Sterman, 2000) and Bounded Rationality (Simon, 1957, 1991). The double-loop learning concept has been used in the model development phase, where the iterative feedback structure has contributed to changes in the pre-conceived mental sub-models of the people in the service department, and has increased the understanding of the department's dynamic structure and mental behavior. The use of bounded rationality has helped us simplify the choice variables by focusing on only three input KPIs in the design stage, as described earlier (user recommendation of the department, workload and competence fit). The principles of system dynamics provide a rigorous approach to strategy that is grounded in solid evidence from the situation concerned (Warren, 2005) and can be used for many decision problems in management accounting.

The paper also shows that the SDM approach can overcome some of the inherent limitations of using only unidirectional cause-effect relationships and statistical methods (e.g. multiple regressions) and the static structure over time, meaning using feed-forward and feedback for testing different strategy scenarios. Compared to earlier research (e.g. Akkermans and Oorschot, 2005), this paper has included some financial KPIs and thereby emphasized the value of the learning outcome from the project modeling project.

Given the huge amount of 'big data' available both from internal but also from external databases (e.g. Orbit or Bloomberg) today, researchers in the field of management accounting will be able to construct and test different types of more generic cost management models (in the same way as within finance), but practical managers will also have the potential to quickly assess their strategic options and opportunities and to design their first version of a dynamic BSC.

Therefore, great opportunities exist now for designers of management accounting programs to develop the right skills for CFO and controllers to be able to handle these challenges in future, by combining 'business analytics' with management accounting and control as also noted by Kaplan.[3]

Reference

ACCA (2009). The CFO's New Environment, A Report Prepared by CFO Research Services in Collaboration with ACCA, UK.

Accenture (2013). Analytics in Action: Breakthroughs and Barriers on the Journey to ROI.

Ahrens, T., and Chapman, C. (2007). Management Accounting as Practice. *Accounting, Organizations and Society*, 32(1–2): 1–27.

Akkermans, H.A. and van Oorschot, K.E. (2005). Relevance Assumed: A Case Study of Balanced Scorecard Development Using System Dynamics, *Journal of the Operational Research Society*, 56(8): 931–941.

Anthony, R.N. (1965). *Planning and Control Systems: A Framework for Analysis*, Harvard University, USA.

Barlas, Y. (1996). Formal Aspects of Model Validity and Validation in System Dynamics', *System Dynamics Review*, Fall, 12(3): 183–210

Brinberg, D., McGrath, J.E. (1985). *Validity and the Research Process*, Sage Publications USA

Bukh, P.N. and Malmi, T. (2005). In Jönsson, S. and Mouritsen, J. (Ed.): *Re-examining the Cause-and-Effect Principle of the Balanced Scorecard' in 'Accounting in Scandinavian – The Northern Light*s, Liber & Copenhagen Business School Press, Denmark.

Capelo, C and Dias, F.F. (2009). A System Dynamics-Based Simulation Experiment for Testing Mental Model and Performance Effects of Using the Balanced Scorecard, *System Dynamics Review,* 23(1): 1-34.

CIMA (2008). Improving Decision Making in Organizations Un-Locking Business Intelligence.

CIMA (2009). Management Accounting Tools for Today and Tomorrow, UK.

Collier, P.E., Berry, A.J., Burke, G.T. (2007). Risk and Management accounting, 2007, CIMA Publishing.

Davenport, T.H. (2006). Competing on Analytics', *Harvard Business Review*, January, 84 (1): 99–108.

Davenport, T.H. and Harris, J.G. (2007). *Competing on Analytics*, Harvard Business School Press, USA.

De Geus, A.P, (1992). Modeling to Predict or to Learn, *European Journal of Operational Research,* 59 (1): 1-6.

De Geuser, F., Mooraj, S. and Oyon, D. (2009). Does the Balanced Scorecard add Value? Empirical Evidence on its Effect on Performance', *European Accounting Review*, 18(1): 93–122.

Eberlein, R.K. and Peterson, D.W. (1992). Understanding Models with Vensim, *European Journal of Operational Research*, 59(1): 216–219.

Eckerson, W. (2011). Big-data analytics: Profiling the use of Analytical Platforms in User Organizations, Business Applications and Architecture Group, Tech Target.

Forrester, J.W. (1958). Industrial Dynamics Major Breakthrough for Decision Makers', *Harvard Business Review*, 36(4): 37–66.

Forrester, J.W. (1991). System Dynamics and the Lessons of 35 Years, WP MIT.

Flamholtz, E. (1979). Toward a Psycho-Technical Systems Paradigm of Organizational Measurement, *Decision Sciences* (January): 71-84.

Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica* 37 (3): 424–38.

Hirsch, M.L. (2002). *Advanced Management Accounting*, 2ed. Thomson, Singapore.

IBM (2011). The power of Analytics for Public Sector, IBM Global business services Government Executive Summary.

IMA (2008). Definition of Management Accounting, Institute of Management Accountants, Montvale, NJ, USA.

IMA (2008). Meet Bob Kaplan, March I, *Strategic Finance*, 19-21.

Kaplan, R.S. (1998). Innovation Action Research: Creating New Management Theory and Practices, *Journal of Management Accounting Research*, 10: 89-117.

Kaplan, R.S. (2011). Accounting Scholarship that Advances Professional Knowledge and Practice, *The Accounting Review* 86 (2): 367–383.

Kaplan, R.S. (2012). The Balanced Scorecard: Comments on Balanced Scorecard Commentaries, *Journal of Accounting & Organizational Change,* 8(4): 539-545.

Kaplan, R.S. and Norton, D. P. (1992). The Balanced Scorecard – Measures that Drive Performance, *Harvard Business Review,* 70(1): 71-79.

Kaplan, R.S. and Norton, D.P. (1996). *The Balanced Scorecard – Translating Strategy into Action*, Harvard Business School Press, Boston, USA.

Kaplan, R.S. and Norton, D.P. (2007). Using the Balanced Scorecard as a Strategic Management System, *Harvard Business Review,* July-August 85(7): 150-161.

Kaplan, R.S. and Norton, D.P. (2008). Mastering the Management System, *Harvard Business Review,* 86(1): 62-77.

Kasanen, E., Lukka, K., and Siitonen, A. (1993). The Constructive Approach in Management Accounting Research, *Journal of Management Accounting Research*, 5:243-264.

Lane, D.C. (1992). Modeling as Learning: A Consultancy Methodology for Enhancing Learning in Management Teams, *European Journal of Operational Research*, 59 (1): 64-84.

Liberatore, M.J., Luo, W. (2010). The Analytics Movement: Implications for Operations Research*, Interfaces*, 40(4): 313–324.

Lyneis, J.M. (1980). *Corporate Planning and Policy Design: A System Dynamics Approach*. MIT Press: Cambridge, MA (now distributed by Pugh-Roberts Associates, 41 Linskey Way, Cambridge, MA 02142).

Malina, M.A. and Selto, F.H. (2001). Communicating and Controlling Strategy: An Empirical Study of the Effectiveness of the Balanced Scorecard*, Journal of Management Accounting Research*, 13: 47-90.

Malina, M.A., Nørreklit, H. and Selto, F.H. (2007). Relations Among Measures, Climate of Control, and Performance Measurement Models, *Contemporary Accounting Research*, 24(3): 935–982.

McKinsey (2011). Competing through data: Three experts offer their game plans.

Meadows, D.H. (1980). *The Unavoidable A Priori, in Elements of the System Dynamics Method*, (ed. J. Randers), MIT USA.

Merchant, K.A. (2012). Making Management Accounting Research more Useful, *Pacific Accounting Review*, 24(3): 334-356.

Nielsen, S. and Nielsen, E.H. (2008). System Dynamics Modeling for a Balanced Scorecard: Computing the Influence of Skills, Customers, and Work in Process on the Return on Capital Employed', *Management Research News*, 31(3): 169–188.

Nielsen, S., Nielsen, E.H., and Skydstofte, L. (2010). Konkurrence på det Analytiske eller 'Competing on Analytics' – Hvordan kan Performance Management, BSC og System Dynamik Kombineres for at sikre Optimal Gennemførelse af Virksomhedens Vision og Strategi?, Økonomistyring & Informatik, 25(5): 413-436.

Nielsen, S. and Sørensen, R. (2004). Motives, Diffusion and Utilisation of the Balanced Scorecard in Denmark, *International Journal of Accounting, Auditing and Performance Evaluation*, 1(1): 103–124.

Otley, D. 2001. Extending the Boundaries of Management Accounting Research: Developing Systems for Performance Management, *British Accounting Review*, 33(3): 243–261

Otley, D. (2003). Management Control and Performance Management: Whence and Whither? *The British Accounting Review*, 35(4): 309-326.

Richardson, G.P., and Anderson, D.F. (1995). Teamwork in Group model building, *System Dynamics Review,* 11 (2), 113-137.

SAS-Institute (2011). Analytics for Government The Smart Way to Cut Costs, Optimise Performance and Deliver Reform in the Public Sector.

Salterio, S. (2012). Balancing the Scorecard Through Academic Accounting Research: Opportunity Lost? *Journal of Accounting & Organizational Change,* 8(4): 458-474.

Sandalgaard, N. and Bukh, P.N. (2008). Danske Virksomheders Budgetpraksis: Udbredelse og Anvendelse, *Working Paper Series Department of Business Studies* No. 18.

Senge, P. (1990). *The Fifth Discipline: The Art & Practice of The Learning Organization*, Random House, Business, New York.

Senge, P and Forrester, J.W. (1980). Test for Building Confidence in System Dynamics Models, TIMS Studies in the Management Science, 14: 209-228.

Silvi, M.R., and Möller, K. (2013). A Framework for Business Analytics in Performance Management, *International Journal of Productivity and Performance Management,* 62(1): 110-122.

Simon, H.A. (1957). *Models of Man*, New York 1957.

Simon, H.A. (1990). Prediction and Prescription in System Modeling', *Operation Research*, 38(1): 7–14.

Simon, H.A. (1991). Bounded Rationality and Organizational Learning', *Organization Science*, 2(1): 125–134.

Speckbacher, G., Bischof, J. and Pfeiffer, T. (2003). A Descriptive Analysis on the Implementation of Balanced Scorecards in German Speaking Countries', *Management Accounting Research*, 14(4): 361–387.

Sterman, J.D. (2000). *Business Dynamics. System Thinking and Modeling for a Complex World*, McGraw-Hill Higher Education, Boston USA.

Sterman, J.D., Repenning, N.P., Kofman F. (1997). Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement, *Management Science*, 43(4): 503-521.

Wallenburg, C.M. and Weber, J. (2006). Ursache-Wirkungsbeziehungen der Balanced Scorecard – Empirische Erkenntnisse zu ihrer Existenz, *Controlling & Management,* 50(4): 245-256.

Warren, K. (2004. Why has Feedback Systems Thinking Struggled to Influence Strategy and Policy Formulation? Suggestive Evidence, Explanations and Solutions', *System Research and Behavioral Science*, 21(3): 331–347.

Warren, K. (2005). Improving Strategic Management with the Fundamental Principles of System Dynamics', *System Dynamics Review*, 21(4): 329–350.

Warren, K. (2008). *Strategic Management Dynamics*, John Wiley & Sons, Ltd. UK.

Vennix, J.A.M. (1999). Group Model-building: Tackling Messy Problems, *System Dynamics Review*, 15(4): 379-401

Wiersma, E. (2009). For Which Purposes do Managers use Balanced Scorecards? An Empirical Study', *Management Accounting Research*, 20(4): 239–251.

Notes

1. The authors would like to thank two anonymous reviewers for their comments and suggestions. The authors are also grateful to the company in our case study for allowing us to use their materials and other documentation and for their support and comments on this research.
2. It should be noted that the above model consists of far more loops than the eight loops shown in figure 2. The eight loops shown here are considered to be of primary importance for the dynamics of the model.
3. This challenge has already been taken up by CIMA (the Chartered Institute of Management Accountants) in UK (see e.g.: http://www.cimaglobal.com/Events-and-cpd-courses/Events/Mastercourses/Strategic-Management/Business-analytics/).

This article explores two issues of
among employees in the public sec
tors and stress. It begins by conside
of performing a group-based compa
present survey data from three grou
context-sensitive approach is advoc
lowing for combinations of contribu
stress in one group may be best exp
model that looks at workload, clear
control over the work situation, this
whole story. The group emerging as
stress was found to comprise emplo
freedom in their work, are highly en
and who set high standards for themsel
that performance indicators did not he
about what was expected of them at w

How Performance Management may lead to Stress

A group-level analysis of performance measurement and employees' stress

Peter Dahler-Larsen and Signe Pihl-Thingvad

Abstract

This article explores two issues of central importance among employees in the public sector: performance indicators and stress. It begins by considering the pros and cons of performing a group-based comparison, then goes on to present survey data from three groups of employees. A context-sensitive approach is advocated for the analysis, allowing for combinations of contributing factors. While stress in one group may be best explained using a classical model that looks at workload, clear indicators, and lack of control over the work situation, this does not tell the whole story. The group emerging as the most vulnerable to stress was found to comprise employees who experience freedom in their work, are highly engaged in their activities and who set high standards for themselves. This group felt that performance indicators did not help to create clarity about what was expected of them at work.

1. Introduction

Over the last couple of years, a prominent subject of debate particularly in the public sector has been performance management and its effects on employees. According to critics, it is difficult to convert core public services into quantitative operational indicators. As a result, there is a risk that performance indicators with low validity become goals in themselves, potentially distorting the work of public professionals (van Thiel & Leeuw 2002; Radin 2006). The various measurements are seen as a sign of mistrust towards public employees (which could promote further mistrust); and it is argued that public employees spend too much time documenting their performance – to the point that their focus is taken away from direct services to citizens (Radin 2006; Van Dooren *et al.* 2010).

On the other hand, it is difficult to imagine modern management *not* being based on systematic data about performance. Advocates of performance management argue that performance information could even prove beneficial for employees (Fletcher & Williams 1996; Van Dooren *et al.* 2010; Walker *et al.* 2012). Well-defined expectations about performance could potentially remove confusion about what is expected of employees and pave the way for greater autonomy as they

decide *how* to achieve the required results (Moynihan 2008). The documentation of results could also play a role in organizational development initiatives (O'Toole & Meier 2011; Boyne 2012).

Performance management has been widely discussed as an ideal typical management approach. However, knowledge is still limited about how much of their daily work groups of employees spend on performance measurement, just as we have very little knowledge about the mechanisms and contextual conditions influencing how performance measurements work in a specific situation.

It is a complex field. Performance *indicators* are criticized on the basis that numbers themselves don't really play a role apart from the organizational context in which they function. At the same time, performance measurement systems change frequently; it is also difficult to isolate their effects from the way management and employees perceive and use them. Indicators are difficult to separate from perceptions of the purpose and value of a given public activity too. The public service may be the subject of disagreement politically, for example. These are just some of the circumstances that make it difficult or meaningless to look at performance measurement as a general and abstract phenomenon (Dahler-Larsen 2014a).

The aim of this article is to help increase knowledge on two fronts concerning performance management in the Danish public sector:

- First, our aim is to assess how much time and mental energy is being allocated to performance measurement documentation, according to public employees.
- Secondly, we want to explore the relationships between performance management and employee stress – by determining whether the group who seems to be experiencing the most stress is also that for which performance measurements appear to be the most extensive. It must be taken into account how each group perceives performance measurements and their impact. Our aim, then, is to develop hypotheses about the different conditions of the groups in order to understand how performance management works in different contexts.

Analyses take place at a social system level rather than at an individual level (Hofstede *et al.* 1993:487). To this end we use average scores enabling us to compare the groups for discussion in this article.

2. Studying the psychosocial work environment

Traditionally, research into the work environment has been largely concerned with identifying objective factors that cause stress. In the *physiological* work environment such factors have a *physical* presence. The rise in knowledge-based work in recent years, however, with its emphasis on flexibility and individualization in organizational tasks, structures and cultures, has changed the focus to the *psychosocial* work environment. This has created challenges for researchers.

For example, it is less clear how to identify and measure evidence of psychosocial strain – because such factors are less objectively measurable, and depend on how they are *perceived* culturally, collegially and individually. From the individual's perspective, management, organization and collegial relations constitute important parts of the psychosocial work environment. It is *people*, then, that to a large extent make up and influence someone's psychosocial work environment. This demands a new approach to research into the work environment, in which the traditional spatial and physiological paradigm of »the individual« exposed to »strains« in »the environment« is replaced by other more systemic ways of thinking (Allvin & Aronsson 2003; Sørensen *et al.* 2012; Pihl-Thingvad 2014a).

Increased individualism within the workforce poses a further challenge. As greater diversity and flexibility is introduced into the way people work, it becomes more difficult to view working conditions as something generic, structured and predictable. Rather, individuals are seen to have more influence over and responsibility for their own levels of stress. This has led to a rise in stress-avoidance or stress-busting initiatives that focus on the individual's mind and inner wellbeing. Politically, the emphasis then moves away from common, structural solutions, as some researchers have observed (Bovbjerg 2011: 66). Emphasis of the work *environment* in research terminology remains popular however (Allvin & Aronsson 2003), despite the emphasis having shifted from the physiological and tangible to the psychosocial. In this spirit, our project seeks to identify organizational factors and work conditions that may influence and cause stress for particular types of work groups.

Two models in particular have dominated stress research to date. Karasek's (1979) demand-control (DC) model sees stress as the result of a situation in which the employees are met with high demands but lack control over the means through which they are able to deliver on these demands. This situation can be improved with social support from management or colleagues (Johnson and Hall 1988; Karasek og Theorell 1990). Siegrist's (Siegrist 1996; 2012; Siegrist *et al.* 2004) effort-reward-imbalance (ERI) model, meanwhile, sees stress as the result of an imbalance between the employees' efforts and the rewards they receive. The imbalance is worsened if the employee is overcommitted to the job. This model is subject to structural conditions at the workplace and the labor market, in that job insecurity may be considered a missing reward for example. However, the model is general and we know very little about what is considered »effort« and »reward« to an individual in a specific situation. Similarly the optimum amount of commitment may be difficult to determine until it's too late.

Empirical research around either of these two models typically involves survey-based studies. These studies – like our own – rely on subjective reporting of *perceived* strains. Additionally, psychosocial work environment research is often subject to endogeneity – ie factors affect each other mutually, making it difficult to separate cause and effect. For example, certain factors might cause stress, and once a person has reached a stressed state they will experience and report these

factors as seeming much more stressful. These complexities aside, it is still relevant to identify *what* is considered stressful at a group level.

3. Important notes about methodology

The results discussed below are derived from the study, *The Significance of Performance Indicators for the Psychosocial Work Environment of Public Employees*, financed by The Danish Working Environment Research Fund. The research was conducted as a mixed-methods study. A more thoroughly discussion of the methods can be found in Dahler-Larsen (2014b).

The results in this article are based primarily on the input of 2941 survey respondents from three groups – language teachers, upper secondary school teachers, and employment consultants. The three groups were chosen on the basis that the represented employees carry out typical modern welfare state tasks *and* are potentially candidates for stress. Previous studies have shown that teachers especially are at risk of stress, as are people who are regularly confronted with the personal and emotional problems of the people they deal with every day (NFA: 2012; Pugliesi 1999) – for example those working with the unemployed, or with students undergoing tests and exams. The three groups were also chosen because they are exposed to diverse performance measurement systems.

Across our three groups we expected that interesting and controversial issues would surface in relation to performance indicators. At job centers, employees experience intense public control in the form of statistical indicators and benchmarking between centers. The employment arena receives much public and political attention; statistically it is one of the most scrutinized policy areas of the entire public sector in Denmark. Indicators showing employment progress, and in getting particular groups of unemployed people into work, are published on a website – www.jobindsats.dk – making it possible to compare the relative performance of job centers on an ongoing basis.

The employment sector is unusual compared to many other public sector departments in that consultants may have come from a range of different educational and employment backgrounds – for example social worker, teacher or other profession. In the teaching profession, by contrast, employees are likely to have followed a more predictable path to their careers.

In recent years, upper secondary schools have experienced a strong control-based reorientation of an otherwise classic professionalized area. These schools are financially dependent on students' enrolment and course-completion rates – ie on the performance of students, which in turn reflects their classroom experience and the quality of teaching. To this end, new assessment mechanisms have been applied in the form of student satisfaction surveys and publication of average grade performance. The professional credentials of upper secondary teachers come from academic degrees, although individual disciplines may range from humanities, social science and business to the natural sciences.

For language instructors teaching Danish to immigrants, the financial situation of their schools is almost exclusively based on the number of students completing each module against an elaborate national system of module tests corresponding to approximately 20 levels of competence in Danish. Language teachers may have different educational backgrounds, typically having completed supplementary education after being teachers, academics or something else.

Our survey went out to all language teachers providing services to immigrants in Denmark. In the case of the upper secondary school teachers, we selected a finite number of schools to target. Similarly for the employment consultants, we selected a fixed number of job centers for our research. For further details on our research methods, see Dahler-Larsen and Pihl-Thingvad (2014b).

3.1. Operationalization of our main variables

»Performance Measurements« are difficult to define – both theoretically, and practically. While any practical operationalization of our theoretical term is debatable, we found the Danish description »resultatmålinger« more fitting for our empirical work. Even so, there was the problem that specific actors in a given context do not identify a certain empirical phenomenon as »resultatmåling«, even though a theorist would do – or vice versa.

We chose not to ask only about specific forms of performance measurement in each organization. This is because we could not predict all practical use of language about such phenomena in advance, and we could not discount the possibility that many other forms would be present in each individual organization besides those which we as survey designers were able to identify. Although we acquainted ourselves with the national and regional performance measurement systems relevant for each group (through document studies and interviews), it would have been impossible for us to have predicted local systems present within each organization. At the same time we wanted a generic, common term so that wording alone would not create systematic differences between the respondents. In practice, we offered the respondents the following guidelines in the survey:

By »performance measurement« (in Danish »resultatmåling«) we mean a numerical calculation of the work by each individual employee – or group, or section, or the entire organization – that is done continually or periodically, and which can be documented or reported upwards in the organization. We term this performance measurement, irrespective of whether the registration takes place as a part of daily work, in separate documents, or electronically.

Performance measurements include, for example, the number of cases processed, the number of consultations or specific results for citizens, users, clients or patients. Performance measurements might also include time spent on specific tasks.

With this broad definition in mind, please provide your opinion on performance measurements in your organization.

We also gave examples of performance indicators as inspiration for each group – such as tests for the language teachers. While this compromise between theoretical definition and everyday language constituted our operational definition of »resultatmåling«, we will refer to both »performance indicators« and »performance measurement« in the text that follows.

The concept of stress is much debated in the research, as stress is considered both something emotional, a mental condition, and an external strain or disease (Bovbjerg 2011; Buch *et al.* 2009:23; Bloch 2001:34; Agervold 1998). Thus, there is both a lack of clarity and a certain complexity surrounding the concept of stress (Allvin *et al.* 2006:135). This article aligns its definition with the traditional perception of stress in the classic work environment research. »Work-related stress occurs when the demands of the working environment exceed the employee's ability to fulfill (or cope with) these« (Research on Work-related Stress. Det Europæiske Arbejdsmiljøagentur, 1999. Luxembourg 2000). For our purposes, then, stress is not considered a disease but a general response to specific incidents or working conditions, which over a longer period of time might result in both a physiological and psychosocial reaction in an individual (see e.g. Allvin *et al.* 2006:146).

The stress level of employees was measured by a number of standardized questions from COPSOQ (The Copenhagen Psychosocial Questionnaire II medium size). The following questions were used to measure stress:

These questions relate to how you have felt over the last 4 weeks.

- *How often have you had problems relaxing?*
- *How often have you felt irritable?*
- *How often have you felt tense?*
- *How often have you felt stressed?*

(Response options: All the time, A large part of the time, Part of the time, A small part of the time, Not at all).

The responses were indexed and converted to a scale ranging from 1 to 100. The scale has a Cronbach's alpha of 0.877.

3.2 The level of our analyses

In the following analyses we compare groups, not individuals – specifically the average statistical scores of the three groups. We do not ignore that a certain level of stress in a group may be serious for the individuals who perceive it, even if the average level of stress appears relatively low in the inter-group comparison. For additional analyses of the relationship between performance measurements and stress on an individual, within-group level, see Dahler-Larsen and Pihl-Thingvad (2014a).

The groups manifest many relative differences. If we examine one difference at a time, we risk overestimating or misrepresenting this factor's influence. Yet we do not offer results based on statistical correlations. It is important to emphasize that we only compare whole groups, and that we offer *hypotheses* to explain the differences in stress levels between them. Our reasoning is that a given combination of factors in a group can lead to further findings, and sometimes a more comprehensive understanding, than can be achieved by focusing on one factor at a time.

Furthermore, our group-based analyses do not allow us to exclude factors that might induce stress at an individual level. Variables to explain differences in stress between two individuals within a group (for example gender, age etc.) might be different from variables that explain differences between groups. The same variables may play different roles at different levels of analysis (Hofstede *et al.* 1993: 487). There is scope then for us – and for other researchers – to find alternative or additional explanatory factors and mechanisms in subsequent analyses.

The comparison between groups is appropriate however, as it encourages consideration of structural conditions and contextual circumstances that would remain invisible when comparing individuals *within* a group. It is also appropriate to look for combinations of factors that characterize groups, rather than isolating individual variables. Most stress models operate with combinations of factors. Here, we will develop hypotheses about which combinations of factors might help us understand why the three groups do not exhibit the same levels of stress. Of course the three groups are not definitely representative of all public employees. However it is still valuable to analyze the groups in relation to the two aims of this article described above.

4. Results

We have already acknowledged that performance indicators rarely work in isolation. With this in mind, we will start by describing the three groups with respect to their perceived stress levels, followed by their general work situations – to provide background knowledge about how performance measurement is experienced in each of the three contexts. On that basis, we will finally look at how performance indicators may contribute to stress. The logic of the text follows the logic of table 1.

Table 1. Stress level, working conditions, and various aspects of performance indicators in three groups of employees

	Language teachers	Upper secondary school teachers	Employment consultants
STRESS LEVEL			
Stress level (0-100)	33.5	34.8	31.4*
WORKING CONDITIONS			
Most stress-inducing factor of the work environment	»My own work standards«	»My own work standards«	»High work pace«
Do you decide the content of your work yourself?	3.21	3.28	2.55*
Do you have to deal with other people's personal problems in your work?	3.76	3.69	4.21*
I'm willing to work a bit beyond what I'm paid for if it helps achieve professional quality in my work	3.55*	3.73*	3.25*
PERFORMANCE MEASUREMENTS			
How often do you think about performance indicators?	3.29	2.84*	3.38
Performance indicators are stressful	2.72	2.58*	2.71
How often does management react negatively to performance measurements?	2.64*	2.79*	3.11*
Performance indicators do not cover the core elements of my work	3.56	3.60	3.71*
Performance indicators lead to a one-sided focus	3.76*	3.48*	3.59*
New forms of performance measurement are being implemented constantly	2.87*	3.21*	3.43*
Performance measurement implies a lack of trust in my ability to do my work properly	2.59	2.90*	2.64
My professional group is able to influence how performance measurement is carried out at my workplace	2.55	2.49	2.71*
Performance measurement supports development initiatives at work	2.68	2.63	2.86*
Performance measurement creates clarity about what we are required to do	3.33*	2.88*	3.62*

Note: Unless indicated otherwise, the average scores are measured on a Likert scale, where 5 equals »strongly agree« or »very often«. The Stress measure is based on an index ranging from 0-100.
* This group differs significantly from the other two in a t-test.

The table shows the average stress level in each group on a scale of 1 to 100. It is significantly lower among employment consultants compared to the other groups, whereas language teachers and upper secondary school teachers have a higher stress level but do not differ significantly from each other.

The next section of the table shows some key characteristics of the working conditions of the groups. Respondents had the opportunity to assess the degree to which a number of different factors in their work were considered stressful. Upper secondary school teachers and language teachers indicated that their own

standards for performance – ie. their own demands of themselves – were the most stressful factor in their work. For employment consultants, pace of work was deemed to be the most stressful factor.

Employment consultants differ from the other two groups in that they have far less influence on the content of their own work. As part of a larger organization bound by a number of managerial targets and legislative regulations, they do not enjoy the freedom of planning their own daily work as is the case with a teacher. All other things being equal, a traditional stress model such as the demand-control model predicts that high demands combined with the lack of freedom to meet them will result in stress. Here, however, consultants' overall self-reported stress level is lower than that exhibited by the other groups.

Exposure to other people's personal problems has also been cited as a source of stress in many public-sector jobs (Sørensen *et al.* 2007). All three groups in this study have a relatively high score on this variable. Language teachers often find themselves faced with the social and personal problems of immigrants. Upper secondary teachers may have to cope with the insecurities of young people – those on the verge of dropping out of school, or worried about not achieving the grades they need to pursue the higher education path or career of their dreams. Meanwhile employment consultants report, to a (significantly) higher degree, that they too have to relate to clients' personal problems. Even so, as a group they are *less* stressed than the other groups.

As a measure of the groups' levels of professional commitment, we asked respondents to rate how far they are willing to work more than paid for in the interests of achieving professional quality in their work. Here, the upper secondary school teachers scored highest, followed by the language teachers.

Summing up; a relatively high pace of work, lack of freedom to plan and organize their own work, and exposure to the personal problems of others were found to characterize the *least* stressed of our three employee groups, rather than the most stressed as might be expected for example in the DC model. The most stressed groups, on the other hand, was characterized by high personal standards and high levels of professional commitment. This observation is consistent with findings in recent stress research (Buch *et al.* 2009; Allvin *et al.* 2006) which has found that, among knowledge workers in particular, the combination of a high degree of autonomy and high personal standards can trigger stress. If high work standards experienced as stress factors represent over-commitment, these findings are also consistent with the ERI model.

If this short summary may suffice as an overview of some of the general factors in the work situation of each of the groups, we can go on to look more specifically at performance indicators.

Upper secondary school teachers devote considerably less time thinking about performance measurements than the two other groups. They also experience stress from performance measurement to a significantly lower degree than the two other groups. On the other hand, for language teachers and employment consultants, institutionalized performance measurements are very prominent in daily practice, both in terms of how much they think about indicators and how stressful they perceive these indicators to be.

We invited all respondents to rate a wide variety of potential stressful factors. The highest ranking factor is work pace (for employment consultants) and personal work standards (among both groups of teachers), as shown in the table. However, given the formulation of the question, all groups place performance measurement after such factors as work pace and their own standards as triggers of stress (the full analysis is presented in Dahler-Larsen & Pihl-Thingvad 2014a). *In other words, there is no support for the claim that performance measurement in itself should generally constitute a predominant problem compared to other stress factors.* If performance indicators lead to stress, our interest should be directed more towards their interaction with specific contexts and mechanisms.

Management most often reacts negatively to performance measurement among employment consultants, less so with upper secondary school teachers and least with language teachers.

That performance measurement does not cover core work tasks, and leads to a one-sided focus on the points being measured, are views reported by employment consultants and language teachers respectively. In neither case is this view particularly pronounced for upper secondary school teachers, however.

The frequency with which new performance measurements are introduced is also a point of discussion. Here employment consultants score highest, and the upper secondary school teachers in between, whereas language teachers have operated with the same basic module system for some years and thus report a lower frequency of change.

In seeking to determine whether performance indicators might be instrumental in explaining stress, we must take into account the following at this point. The most stressed groups do not experience the most common negative management reactions to performance measurement; nor do they succumb to the most negative distorting effects of performance measurements, nor the most frequent changes to performance measurement systems.

On the other hand, we need to consider the following variables. Upper secondary school teachers believe that performance measurements indicate a *lack of trust* – to a larger extent than the other groups. This could be explained logically in terms of self-awareness resulting from their educational background, their high profes-

sional involvement, and their extensive history as relatively autonomous professionals (at least until the recent upper secondary school reform).

The question as to whether performance measurements help to create *clarity* reveals the largest differences between the groups in the entire table. Upper secondary school teachers register the lowest score. In their world, there are several performance indicators at play (pay-per-pupil systems, drop-out rates, grades and student satisfaction surveys). In addition, the conflict between performance measurements and the employee's own professional values are strongest among upper secondary school teachers. They all have an academic background and, of our three groups, have the highest professional involvement and the most intense conflict between their own personal standards and their actual work practice (results are reported in Dahler-Larsen & Pihl-Thingvad 2014a). Given this background, they might also have the highest degree of resistance to the idea that performance indicators define a legitimate aim of their work. Among employment consultants, on the other hand, we find that performance measurements not only create more clarity over what is expected of them, but also that these indicators lead to development initiatives at the workplace and that employees have a say in how performance measurements are used. These three positive conditions of performance measurements seem to suggest – at least in terms of this group comparison – that the practice of performance management has at least some advantageous functions for employment consultants.

If performance measurements are to help explain that the two groups of teachers are the most stressed of the three groups in the study, the mechanisms are *not* negative management reactions or performance measurements' distortion of work. The most likely mechanisms, we hypothesize based on the data, are rather connected to the feeling of distrust, the sense of a lack of influence in the professional group, and the fact that performance measurements do not appear to create clarity nor help to promote development initiatives. All this can be understood in the light of these employees' own sense of professionalism. A key role in this story is also played by the teachers' own work standards, ie. their personal expectations of themselves and of the work they do. A source of tension may lie in a discrepancy between their own standards and the lack of perceived meaningfulness and clarity in the performance indicators. For a further discussion of the relationship between employees' own work standards and performance measurements within groups, see Pihl-Thingvad (2014b). In any case, this combination of factors is our best candidate for an explanation.

5. Discussion and conclusion

The three groups of employees in our study did not experience performance measurements as the most stressful factor. Typically, »own standards« and »work pace« are perceived as more stressful.

The group that spends most time thinking about performance measurements daily is not the most stressed group. Nor is the lack of the freedom in determin-

ing tasks, or the need to confront the personal problems of others, particularly pronounced in the most stressed group. If performance measurements contribute to stress in the most affected groups, it is not connected to a high level of negative reaction from management, nor to any distortion of their work due to performance measurement. Quite contrary, for example, upper secondary school teachers, a highly stressed group, is characterized by feelings of distrust, of a lack of professional influence, and a significant sense that performance measurements do not provide clarity around work expectations. The professional involvement and high personal standards of the teachers might contribute to bringing these mechanisms into play. These findings are consistent with the ERI model. *Although you experience freedom in your work, the work might still be stressful when you are very engaged in it, set high standards for yourself (which are then experienced as stressful), and furthermore feel exposed to performance indicators that are seen to be based on distrust and which do not create clarity*. The lack of clarity is probably due to both lack of correlation between several indicators (grades, student satisfaction, and drop out) as well as a mismatch between some of these indicators and the professionalism and self-awareness of the teachers.

The differences observed between the groups should not be used as the basis for arguing that individuals in less stressed groups are not stressed. With the least stressed group, another configuration might contribute to stress (although of course at a lower, overall level for the entire group). Fast work pace, lack of job freedom, clients' personal problems and management's negative reactions to performance measurement results might, for example, lead to stress for employment consultants, just as the frequent introduction of new performance indicators might.

The analysis in this article is tentative. Substantiated hypotheses have been developed on the basis of the table and inspired by theories about sources of stress. A direct and isolated impact of performance measurements upon stress is hardly decisive, as it is lowest in the most stress-affected group(s), and we have other – and better – explanations for the stress being experienced here. This being the case, we have offered *configurations* of conditions that can be used for understanding *how* performance measurements work in different contexts at a group level.

Our findings have both managerial and organizational implications. A general, positive prescription for performance measurement does not seem helpful. An understanding of different combinations of contributors is warranted (see also Radin 2006).

For instance, one configuration includes a preoccupation with performance measurements, negative management reaction to the measures, fast work pace and limited professional involvement, but at the same time clarity of demands. This combination of conditions seems to be associated with a certain level of stress. A higher level of stress is found in another configuration characterized by high levels of professional involvement, high personal standards, and the feeling of dis-

trust and uncertainty connected to performance measurement. Factors such as the experience of distrust, involvement of professional groups in the preparation of how to measure performance, as well as organizational development on the basis of performance measurements, seem to be the obvious focus areas – if the aim is to promote the use of performance indicators managerially and organizationally, without simultaneously increasing stress.

In theoretical terms, we are suggesting a situation-specific and contextually-oriented approach to performance management. While the first situation described above (found among employment consultants), may resemble the work conditions described in the demand-control model, the characteristics of the latter situation (for upper secondary school teachers) lies closer to the ERI model.

Our findings about the ability of performance indicators to create clarity (which differs starkly across different contexts) are interesting. Even though all groups widely agree with the statement that performance indicators contribute to a one-sided focus in their work, the groups differ considerably in their assessment of whether performance measurements create *clarity* around what is required of them. Qualitative data from the overall mixed-methods study (interviews with language teachers) indicates that while the module test system evidently has pros and cons, its ability to create clarity about the demands on employees is ultimately a positive factor. Even though such a performance measurement system seems to colonize the definition of sound professional work, it also removes confusion about what is expected (Dahler-Larsen 2014c).

Increasing clarity in this way, then, would seem to be a sensible management strategy. This might involve removing ambiguity from a multitude of interacting or conflicting systems of performance measurement, to bring managerial messages into line with monitoring systems and to keep the frequency of change in performance measurement systems under control. However, all this is easier said than done and, to an extent, controversial. Ambiguity in performance management systems won't be removed by clearer management signals alone. Many performance measurement systems are defined in ways and on levels – for example a national level – which are beyond the reach of the individual organization or individual manager. Furthermore, uncertainty may be a sign of tension within a particular combination of conditions – including professionalism, self-awareness and the individual's own standards of performance for example. A debate about professional values is often also a debate on how to handle broader democratic issues in society.

An important overarching dilemma emerges too – from the observation that those professional groups which set high standards for themselves are also those that experience performance indicators as a source of distrust and uncertainty. A management dilemma, then, is whether to help these employees lower their own standards, with a view to increasing the acceptability of performance indicators. A more productive option might be to establish a dialogue between management

and employees about how individuals' own standards are best understood and calibrated under particular conditions. Although in a modernized, individualistic culture, personal expectations and standards are perceived to be an issue only for the individual concerned, they are organizationally and sociologically the result of social processes (Pihl-Thingvad 2014b).

Our analyses were designed to stimulate debate. In this article we suggested a situation-specific and contextual approach to performance measurements and stress, since different models are appropriate to understanding different work conditions. Beyond that, we have pointed to the importance of two special issues that deserve attention. One is the role of clarity in performance management, the other the role of the employee´s own standards – neither of which should be considered easy to manage. In our overarching research project (Dahler-Larsen and Pihl-Thingvad 2014a), we explore differences between individuals in our groups, and elaborate further on the mechanisms discussed above.

Literature

Agervold, M. 1998. *Det psykosociale arbejdsmiljø – Fra videnskabelig arbejdsledelse til arbejdsmiljøpsykologi*. Aarhus: Aarhus Universitetsforlag.

Allvin, M., Aronsson, G., Hagström, T., Johansson, G., & Lundberg, U. (2006): *Gränslöst arbete – socialpsykologiska perspektiv på det nya arbetslivet*. Malmö: LIBER.

Allvin, M. & Aronsson, G. 2003. The Future of Work Environment Reforms: Does the Concept of Work Environment Apply within the New Economy? *International Journal of Health Services*, 33(1): 99-111.

Bloch, C. 2001. *Flow og Stress – Stemninger og følelseskultur i hverdagslivet*. Frederiksberg: Samfundslitteratur.

Bovbjerg, K. 2011. Arbejdsmiljø og markedet for stresshåndtering. In K. M. Bovbjerg (Eds.), *Motivation og Mismod*. Aarhus: Aarhus Universitetsforlag.

Boyne, G.A. 2012. Performance Management: does it work? In R.M. Walker, G.A. Boyne, & G.A. Brewer (Eds.): *Public Management and Performance. Research Directions*. Cambridge: Cambridge University Press.

Buch, A., Andersen, V., & Sørensen, O.H. 2009. *Videnarbejde og stress – mellem begejstring og belastning*. Copenhagen: Jurist- og Økonomforbundets Forlag.

Dahler-Larsen, P. (2014a) Performance management og resultatmålinger. Chapter 1 in P. Dahler-Larsen and S. Pihl-Thingvad: *Resultatmålinger og Stress.* Odense: Syddansk Universitetsforlag

Dahler-Larsen, P (2014b) Metode. Chapter 3 in P. Dahler-Larsen and S. Pihl-Thingvad: *Resultatmålinger og Stress.* Odense: Syddansk Universitetsforlag.

Dahler-Larsen, P. (2014c) Klarhed og uklarhed i resultatmålinger: Et processuelt perspektiv. Chapter 12 in P. Dahler-Larsen and S. Pihl-Thingvad: *Resultatmålinger og Stress.* Odense: Syddansk Universitetsforlag.

Dahler-Larsen, P. and Pihl-Thingvad (2014a): *Resultatmålinger og Stress*. Odense: Syddansk Universitetsforlag.

Dahler-Larsen, P. and Pihl-Thingvad (2014b): De tre undersøgte grupper og deres resultatmålingssystemer. Chapter 5 in P. Dahler-Larsen and S. Pihl-Thingvad: *Resultatmålinger og Stress.* Odense: Syddansk Universitetsforlag.

Fletcher, C. & Williams, R. (1996): »Performance Management, Job Satisfaction and Organizational Commitment«, *British Journal of Management, 7*: 169-179.

Hofstede, G., Bond, M.H. & Luk, C.-l. (1993): Individual Perceptions of Organizational Culture: A Methodological Treatise on Levels of Analysis, *Organization Studies*, 14(4): 483-503.

Johnson, J.V. & Hall, E.M. (1988): »Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population«, *American Journal of Public Health,* 78*:* 1336-1342.

Karasek, R.A. 1979. Job demands, job decision latitude, and mental strain: Implications for job redesign, *Administration Science Quarterly*, (24): 285-307.

Karasek, R. & Theorell, T. (1990): *Healthy Work. Stress, Productivity, and the Reconstruction of Working Life*. New York: Basic Books.

Limborg, H.J. 2002. *Den risikable fleksibilitet – på vej mod et nyt arbejdsmiljø*. Copenhagen: Frydenlund.

Moynihan, D. 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington DC: Georgetown University Press.

NFA (2012): Følelsesmæssige krav i arbejdet. København: Det Nationale Forskningscenter for Arbejdsmiljø.

NFA (2012): Tredækkeren II. Hentet fra http://www.arbejdsmiljoforskning.dk/~/media/Spoergeskemaer/copsoq/konstruktion-af-skalaerne-i-3d-ii-28-10-20052.pdf#

O'Toole, L.J. Jr. & Meier, K.J. (2011): *Public Management. Organizations, Governance and Performance*. Cambridge: Cambridge University Press.

Pihl-Thingvad, S. (2014a): Hvad er psykisk arbejdsmiljø? – forskning i det psykosociale arbejdsmiljø. Chapter 2 in Dahler-Larsen, P. and Pihl-Thingvad (2014): *Resultatmålinger og Stress*. Odense: Syddansk Universitetsforlag.

Pihl-Thingvad, S. (2014b): Resultatmålinger og egne krav til præstation. Chapter 11 in Dahler-Larsen, P. and Pihl-Thingvad (2014): *Resultatmålinger og Stress*. Odense: Syddansk Universitetsforlag.

Pugliesi, K (1999): The Consequences of Emotional Labor: Effects on Work Stress, Job Satisfaction, and Weil-Being. *Motivation and Emotion*, Vol. 23, No. 2

Radin, B. (2006): *Challenging the Performance Movement Accountability, Complexity, and Democratic Values*. Washington DC: Georgetown University Press.

Siegrist, J. (1996): »Adverse Health Effects of High-Effort/Low-Reward Conditions«, *Journal of Occupational Health Psychology, 1(1),* 27-41.

Siegrist, J. (2012): Effort-reward imbalance at work – theory, measurement and evidence. Düsseldorf: Department of Medical Sociology, University of Düsseldorf, Germany.

Siegrist, J., Starke, D., Chandola, T., Godin, I., Marmot, M., Niedhammer, I. & Peter, R. 2004. The measurement of effort-reward imbalance at work: European comparisons.

Sørensen, O.H., Buch, A., Christensen, P.H. & Andersen, V. 2007. Indflydelse i vidensarbejdet – kan man få for meget af det gode? *Dansk Tidsskrift for Arbejdsliv*, 9(2): 38-54.

Sørensen, O.H., Hasle, P., Hesselholt, R.R., & Herbøl, K. (2012): *Nordiske forskningsperspektiver på arbejdsmiljø. Mening, indflydelse og samarbejde*. København: Nordisk Ministerråd og Nordisk Råd.

van Dooren, W., Bouckaert, G., & Halligan, J. (2010): *Performance management in the Public Sector*. London & New York: Routledge & Taylor and Francis Group.

van Thiel, S. & Leeuw, F. The Performance Paradox in the Public Sector. *Public Performance and Management Review*, 25(3): 267-81.

Walker, R.M., Boyne, G.A., & Brewer, G.A. (2012): *Public Management and Performance. Research Directions*. Cambridge: Cambridge University Press.

This article investigates differences and
between performance management syst
public services. We offer three theoreti
ons as to why the characteristics of perf
management systems may still converge
policy areas in the public sector with dif
of task complexity amid a lack of formal
ching, government-wide policies. We ad
propositions from a comparative case stu
examines the characteristics of perform
nagement systems across social service
care) in contrast to those employed in te
vices (park services) in Denmark. Contr
pectations of divergence according to d
task complexity, the characteristics of
management systems in the two policy
served to converge. We proposes that
has occurred due to 1) similarities in
reforms; 2) institutional pressures
mentarity between political needs
needs. Finally, we discuss how
to existing models of the

Three propositions about why characteristics of performance management systems converge across public sector services with different levels of task complexity

Bente Bjørnholt, Andrej Christian Lindholst and Jeppe Agger Nielsen

Abstract

This article investigates differences and similarities between performance management systems across public services. We offer three theoretical propositions as to why the characteristics of performance management systems may still converge across policy areas in the public sector with different levels of task complexity amid a lack of formal and overarching, government-wide policies. We advance our propositions from a comparative case study which examines the characteristics of performance management systems across social services (elderly care) in contrast to those employed in technical services (park services) in Denmark. Contrary to expectations of divergence according to differences in task complexity, the characteristics of performance management systems in the two policy areas are observed to *con*verge. We proposes that convergence has occurred due to 1) similarities in policy-specific reforms; 2) institutional pressures; and 3) complementarity between political needs and managerial needs. Finally, we discuss how our findings contribute to existing models of the way performance management systems develop in the public sector.

1.0 Introduction

The use of performance measures and management has a long-held tradition in the public sector, and extensive experience has been accumulated (Johnsen, 2005). Since the 1980s especially, the public sector has witnessed a surge in the use of performance management and measurement to ensure efficiency, effectiveness, and accountability (De Bruijn, 2002; Pollitt & Bouckaert, 2004; Taylor, 2011).

Prominent performance management scholars (e.g. Moynihan et al., 2011) have called for a contingency perspective (Chenhall, 2007), according to which optimal performance management systems should be designed and adapted to achieve desired organizational goals and outcomes. Here, 'optimal performance management

system' is contingent on both external and internal organizational factors, including contextual requirements and the relative complexity of the tasks involved in delivering particular services. From this perspective, the characteristics of performance management systems should be expected to vary across given public-sector policy areas and services.

However, it has been noted that the state context also limits variations and design options through formal, government-wide policies which promote the uniform usage of performance management systems across particular services (Radin, 2006). Such moderation may result in the adoption of less-than-optimal performance management systems in some policy areas; therefore caution has been advised in the application of overarching, government-wide policies for uniform performance management models, which risk neglecting the underlying causes of poor performance, and quality of available performance indicators, and could distort incentives in particular policy areas (Curristine et al., 2007, p. 28).

This paper adds to the understanding of performance management systems adoption in the public sector by highlighting further moderating mechanisms in a state context that might explain why performance management systems sometimes convergence across dissimilar policy areas. It explores differences and similarities between performance management systems in different public services. We suggest that convergence in the design (characteristics) of performance management systems, and therefore the potential adoption of less-than-optimal performance management systems, could be the result of a set of complementary factors, in the context of a lack of overarching, government-wide policies. To investigate this claim we compare two different services (elderly care and park services) provided by Danish local government, representing different levels of task complexity, in a national context where no overarching or formal government policies for the uniform use of performance management systems have been implemented.

The comparative case study reveals a tendency for the characteristics of performance management systems within the two services to converge in crucial aspects (purpose, target setting, measurement, and motivational mechanisms). We trace the reasons for this convergence, and advance a set of propositions which add to already established findings about the influx of government-wide policies, and which help explain the convergence of performance management system characteristics in the public sector. We propose that convergence may occur as the result of 1) similarities in policy-specific reforms; 2) institutional pressures; and 3) complementarity between political and managerial needs.

Next, we introduce performance management and its core activities, in order to operationalize the factors investigated in the comparative case study. We then describe our research methods and our reasons for selecting these two particular service areas, before setting out our case study analysis where we compare the characteristics of performance management systems across the two services. Following on from this discussion we propose three theories which we believe

could explain the convergence observed, and discuss how our findings add value to existing models of performance management system development in the public sector.

2.0 Performance management

Performance management systems (Pollitt, 2013) are defined according to a number of characteristics. By 'systems', we mean the way that performance management is incorporated in a strategic cycle of goal-setting, monitoring, and feedback mechanisms (e.g. Bruijn, 2002; Halachmi, 2005). The ideal model for designing performance management systems has come to be associated with the specification of measurable performance indicators, clear performance targets, the transparent monitoring of performance, and tight coupling between performance, budgets, and financial measures. In this model, performance management systems enable decision-makers to set priorities and specify expected performance, and agencies to perform efficiently and in alignment with what is expected of them (de Bruijn, 2002).

Thus, performance management is expected to include the formulation of strategic goal-setting, clarification of the link between a program and its (expected) effects and use of resources, and ongoing monitoring and feedback. Performance management tools can be distinguished by their 1) purposes; 2) performance definitions; 3) methods; and 4) monitoring systems. These four characteristics represent core activities in defining performance management, and are in line with the main considerations when applying performance management systems (e.g., de Bruijn, 2002; Halachmi, 2005; Van Dooren et al. 2010).

'Purposes' are the motives behind the introduction of performance management systems. These are essential as different reasoning or prioritization is likely to result in a different choice of measures (Pollitt, 2013). The systems may serve different purposes, such as learning, control, motivation, budgeting, promoting, or accountability (Brehn, 2003; Van Dooren et al., 2010). Under ideal circumstances, performance systems would have a precise purpose; in practice, however, these are often mixed and unclear (Pollitt, 2013).

Performance definitions include deciding what the object of the performance management system is. Performance management may involve multiple goals and ways to measure them, and different perspectives depending on the stakeholder (Moynihan, 2008). We are primarily interested in identifying who defines the performance indicators, and what the main considerations of causal mechanisms are. Performance management systems produce data which are usually – though not always – quantitative (Pollitt, 2013). Quantitative data methods typically leave less room for actors to influence the standards of assessment and their results during the assessment process. Thus, quantitative measures reflect tighter control (Beavan & Hood, 2006; Bruijn, 2002). Finally, monitoring, which specifies the motivation mechanisms, is essential in characterizing performance systems. If monitoring is linked to the allocation of resources between individuals or organizations, it repre-

sents tight control, providing further performance incentives in conjunction with the measurements (de Bruijn, 2002).

3.0 Reserach Methods

To conduct our research we compared the characteristics of performance management systems used by two Danish service providers run at a local government (municipality) level. Danish municipalities have been at the forefront of implementing performance management systems (Hansen, 2011) and encompass a wide range of services involving task of varying complexity. The two service areas chosen for this study are elderly care (a social service) park services (a technical service), as representing diverse requirements. Successful examples of performance management systems tend to be found within services reflecting a low degree of task complexity, such as technical services (Moynihan et al., 2011), whereas performance management appears to be more problematic in social services where there is typically a higher degree of task complexity (e.g., Van Dooren et al., 2010).

In Denmark, elderly care is one of the welfare state's 'soft' service areas – i.e. involving human care and direct involvement with users. Task complexity is relatively high, and the services attract high public and political attention and tend to be the subject of controversy, operating in a context where professional logic constantly is challenged by managerial logic (Rostgaard, 2012). Elderly care is provided in two formats: *home care*, where clients receive practical or personal help in their homes; and *nursing homes*, where clients live on a permanent basis and receive practical and personal help. Services included both advanced tasks, related to self-help and the rehabilitation of patients in their homes (Nielsen & Andersen, 2006), and a number of standardized tasks such as dusting and cleaning.

By comparison, park services typically receive limited public and political attention and take place within a statutory framework with few or no legal requirements or policies (Lindholst et al., 2014). Provision is characterized by a range of technical tasks, such as tree pruning or grass cutting, as well as managerial tasks, such as planning which services to provide, allocating organizational resources, managing staff, or overseeing contracts (Gustavsson et al., 2005; Lindholst & Bogetoft, 2011). Service standards within the sector are adopted at a decentralized level by professionals with a main focus on the maintenance of green infrastructure rather than a focus on the underlying recreational, functional, or community needs (Lindholst et al., 2015). The management of park services also involves a degree of engagement with users (organized and un-organized). However, the level of engagement, involvement, and direct interaction with users is limited in administrative practice (Jansson & Lindgren, 2012).

As recommended by Yin (2008), our comparative case study builds on multiple data sources. These range from statistics and existing research, to reviews of various performance management systems within the two areas. For *park services*, we draw on several research sources covering marketization and organizational change (Nuppenau, 2009); different approaches for management and contracting

out within park services (Lindholst, 2009) and urban park management and maintenance in general (Bell et al., 2008; Randrup & Persson, 2009); as well as reviews of various performance management systems, including the Yardstick, Parkcheck (Yardstick, 2012), the Nordic Green Space Award (Green Space Award, 2013; Lindholst et al., 2012), and the professional standard for park maintenance in Denmark (Juul et al., 1998). For *elderly care*, we draw on relevant legislation (e.g., the Social Services Act), policy documents and reports (e.g., FOA, 2005; Ministry of Social Affairs, 2007); statistics from Statistics Denmark; and existing research on the nature of major reforms of elderly care in Denmark in the period from the mid-1990s to 2010 (e.g. Bjørnholt, 2006; Nielsen & Andersen, 2006; Rostgaard, 2012).

4.0 Analysis

Our analysis begins by examining the nature of performance management systems within elderly care and park services and their development over time. Based on this analysis, we evaluate the converging and diverging tendencies of performance management systems between the two scenarios.

4.1 Performance Management in Elderly Care

For decades, elderly care service has been associated with images of inefficiency and low status (Rostgaard, 2012), and ranked comparatively low in surveys (Nielsen & Andersen, 2006). In this environment, and especially since the mid-1990s, various stakeholders have called for better management of the sector. Today, satisfaction surveys, benchmarking, and time registration have become integrated aspects of elderly care management. The increased use of performance management systems in elderly care – particularly in home care – has been driven by a number of government reforms, such as Common Language (Hansen & Vedung, 2010), Free Choice (Nielsen & Andersen, 2006; Rostgaard, 2006), and Elderly Care Documentation Project (Ministry of Social Affairs, 2007). Parallel to this development, however, municipalities themselves have developed performance management systems, including municipal-specific targets (Bjørnholt, 2006). Table 1 summarizes the most pronounced performance management systems in elderly care.

The *Common Language Project*, carried out from 1992 to 2004, represents the first major national initiative to improve the management of elderly care. Overall, the opportunities to benchmark elderly care services were boosted, as the Common Language reform established standards and new, compatible ways of describing elderly needs and the services they are entitled to (Hansen & Vedung, 2010). A monitoring and feedback system was never established, as a planned joint national database to benchmark municipalities against each other was never implemented (ibid).

Table 1: Performance Management Systems in Elderly Care

PM system	Purpose(s)	Performance definition	Methods	Monitoring mechanisms
Systems initiated at the national level				
Common Language	Improved management of elderly care	The Local Government Association (KL)	Quantitative. A set of predefined performance indicators	National database never implemented
Free Choice	Accountability and increased marketization through benchmarking against other organizations	Parliament and ministries	Quantitative. A set of predefined performance indicators	Municipalities expected to provide statistics to a national database
Documentation Project	Use benchmarking to enable politicians to hold the municipalities responsible for the results	Parliament and ministries. However, indicators were shaped by what was already measured at municipal level	Quantitative. Focus on result and impact measures	Municipalities required to provide statistics on 23 measures in a national database
Systems initiated at the municipality level				
A variety of performance indicators	Document effective work time with clients (e.g., user time percentage) Information on the quality of services (e.g., the number of different care workers per client)	Politicians and managers at the municipal level, but also shaped by national reforms	Quantitative	Performance targets required to ensure the accurate documentation of the costs
Satisfaction surveys	Information about the quality of services from user/client perspectives	Politicians and managers at the municipal level	Quantitative but also qualitative elements	No sanctions or rewards

In the early 2000s, government reforms continued to support the development of performance measures in elderly care. In particular, the *Free Choice* reform played a crucial role, even though it does not represent a performance management system *per se*. In 2003, a Liberal–Conservative government enacted a requirement for home care services to be open for tenders, the intention being to improve quality, cut costs, and encourage market forces in care provision (Rostgaard, 2012). To support the free-choice model, municipalities are obligated to specify the level of services and costs and make systematic and comparable information available to service users. Accordingly, services are described in great detail, and a certain amount of time is assigned to each task (e.g., 5 minutes for preparing breakfast). Consequently, this reform has fostered a comprehensive performance and benchmark system. First, a more accurate documentation of service costs was required to support a marketization movement, in order to guarantee fair competition between public and private providers. This was followed by a monitoring and benchmark system, as all Danish municipalities are expected to upload their quality standards, prices, and other measures to a national, Internet-based database. These quantitative performance indicators are intended to ensure transparency and comparisons in the area and enable politicians and other stakeholders to control how the Free Choice reform progresses.

More recently, the *Elderly Care Documentation Project* (Ministry of Social Affairs, 2007) has continued the efforts started in the Common Language reform and has succeeded in establishing a national database with a variety of performance indicators. Municipalities are now required to deliver statistics to a national database. As a result, a number of indicators are published on the Statistics Denmark website including seven impact indicators (e.g., user time percentage), as well as 16 background indicators including the number of clients receiving home care. The overall purpose is to provide relevant information to politicians and other key decision-makers. While the Documentation Project was decided and mobilized by the parliament and ministries, it is important to note that many municipalities had already initiated a number of the indicators. The monitoring system is considered to be strict, in that municipalities are obligated to report to Statistics Denmark, although there are no bonus or sanction systems.

While performance management systems have been shaped by government reforms, municipalities are also heavily engaged in the discipline (Bjørnholt, 2006). Below, we provide three examples to illustrate the variety of measures municipalities are using to improve the management of the area and satisfy government requirements. First, the measure of *user time percentage* has been introduced to document effective work time with clients. In practice, care workers are required to report the time spent and services provided at the point of care. The results enable politicians, both at the local and national levels, and other stakeholders to control service costs and service delivery. User time percentage has been included as an indicator in the national Elderly Care Documentation Project recently too, but has long (from before the Free Choice Program) been a key point of measurement in many municipalities. Other measures, such as *the number of different care workers per client,* focus more on service quality. This specific measure is now also included in the national Documentation Project. Measuring *user satisfaction* is mandatory, meanwhile, even though it is up to each municipality to decide the questions in the survey. The purpose here is to provide information about the quality of elderly care from the user/client perspective. At a municipal level, there are also qualitative and municipality-specific measures at play. For instance, these are applied when the users are unable to answer quantitative questionnaires (e.g. those suffering from dementia) or when more in-depth information is needed.

Overall, performance management systems have defined much of the public agenda and the organization of elderly care service provision. A variety of systems and measures have evolved in a recursive relationship between national and municipal-level initiatives and interests. Nonetheless, the many procedures to control time, and to document the use of resources (supported by decree from the central government) have led to intense debate and struggles between stakeholder in the elderly care field (Rostgaard, 2012). Led by the powerful DaneAge Association and trade unions, critics claim that the new management and performance approaches have taken elderly care off course with bureaucracy, the 'tyranny of time', de-motivated employees, and inferior services for the elderly. One of the central arguments has been that care workers currently spend too much time on

administrative tasks instead of caring for clients (e.g., Andersen, 1999; FOA, 2008), reflecting an institutional tension in elderly care between management logic and a more traditional professional logic. Despite this disagreement and criticism, use of performance systems appears to persist. At the same time, counter trends are in evidence: for instance some municipalities have introduced 'zones free of control' and projects aimed at empowering (as opposed to controlling) employees.

4.2 Performance Management in Park Services

As in elderly care, the need to improve efficiency and focus on organizational performance has been an early feature of municipal park management through the implementation of New Public Management (NPM) tools – in the form of quality standards, provider–purchaser models, and frequent use of public procurement and outsourcing (Lindholst, 2009; Lindholst et al., 2015; Nuppenau, 2009). While developments in park services reflect broader trends, there is no evidence of national government initiatives supporting this development. Table 2 provides an overview of three complementary systems which have been specially developed for park services performance management.

Table 2. Performance management systems in park services

Performance management system	Main purpose(s)	Performance definition	Methods	Monitoring mechanisms
Systems initiated at the national level				
None				
Systems initiated at the municipal level				
Standardization	Cost efficiency/rational planning/requirement for management of internal of external arms-length relations	Typically managers and consultants	Quantitative (e.g., grass cutting or tree pruning). Use market to test prices on predefined service levels	Contractual logic, including economic penalties for defaults
Green Space Award	Improve performance/ raise public and political profile (give account)	Managers, research and users, optional for applying for approval of a park for a Green Space Award (GSA)	Qualitative/external inspection and judgment of a park against a predefined set of criteria	Input to planning decisions/resource allocation
Yardstick, Parkcheck	Accountability through benchmarking against other organizations	Managers and consultants, based on a common system	Quantitative/external measurement of a set of predefined performance indicators for organization and services	Input to planning decisions/resource allocation

The initial adoption of performance management was based on a sector-wide system for the *standardization* of maintenance operations (e.g., grass cutting or tree pruning), and for control over maintenance against expected service levels, as defined by a set of common standards. The system was developed in the late 1990s to provide a 'common language' for municipalities. It has been widely implemented by Danish park managers to address the need to improve internal

efficiency and/or outsource services. In contrast to the changes observed in elderly care, this system has been developed from within the sector itself and adopted without government support or statutory requirements issued at a national level. It has been based on a specification of services in technical and quantitative terms fit for the purpose of, for example, benchmarking, competitive pricing, setting maintenance levels, and the rational planning of work. Standardization has been established by defining a range of constituent 'elements' in a green space and a set of performance- and instruction-based specifications and acceptance levels for the quality of associated maintenance tasks. In this case, performance has been managed using a market-based approach, with competitive pricing as the main mechanism for achieving efficiency, and contractually-embedded mechanisms for feedback and the tight enforcement of terms, including economic penalties for defaults (Lindholst, 2009).

In the second half of the 2000s, the sector's standardization system was complemented by other viable performance management systems, partly in response to criticism of the standardization system (Lindholst et al., 2015). Among the newer systems are more holistic approaches at the park level, such as the 'Green Space Award' and organization-wide performance management systems including 'Yardstick, Parkcheck'.

The *Green Space Award (GSA)* is a system for enhancing and benchmarking the quality of park services at an individual park and management team level. The system was developed by a Nordic network including municipalities, research institutions, professional organizations, and interest groups without support or initiatives from central administration or national-level politicians. The development was partly on the result of criticism of the limited scope of the standardization system, which is quantitative in nature, to cover broader structural, functional and recreational qualities. The core of the benchmarking consists of 41 quality criteria grouped within three themes related to 'structure and general aspects', 'functionality and experience', and 'management and organization', which together define excellence in park services. The performance indicators in the GSA system reflect a body of research (e.g. Bell et al., 2008) that identifies a broader range of tangible and intangible public values associated with green spaces and their use. Measurement is carried out in the GSA system through an external audit, performed by a network of trained professionals and laypeople, summarized by a score and written statement. Parks that pass a certain level receive an award that can be used for promotion and recognition. Municipalities themselves choose whether they want to apply for an audit, and there is no tight coupling between the allocation of resources and the result of the audit. The main purpose is to focus internally on excellence and gain external recognition, thereby providing an incentive for the allocation of resources towards parks that are audited. The development and definition of the performance measures has found broad support from decentralized actors including municipalities, research institutions, professional associations, and interest groups within the field. Central political support at a national level has been absent however, although the aim of the GSA fits neatly with the surge

of new performance measures and quality standards for different public services (e.g., in the field of elderly care) that has been coordinated by central state agencies and the Local Government Association.

Yardstick, Parkcheck is an internationally available benchmarking tool developed by parks and facility managers to provide a method for collecting management information and for measuring, evaluating, and driving the performance of public organizations responsible for urban parks and green spaces. Yardstick, Parkcheck establishes a performance management system for benchmarking, which – in theory – demonstrates the relative efficiency of an organization (through comparison either with earlier years or the performance of other organizations). The system has been implemented by a range of major Danish cities since 2010, including Aalborg, Aarhus, and Odense. The rationale lies in a series of key performance indicators (KPIs) measuring performance within both management (e.g. operational excellence or efficiency) and services (e.g. user satisfaction), which can be compared across municipalities, nationally as well as internationally. Information from participating municipalities is updated annually – performance can be evaluated against specific indicators, and areas identified for improvement. Performance is reviewed using a standardized methodology, based on external assessment carried out by a consultancy, which includes both qualitative and quantitative information. Overall, the Yardstick model focuses at an organizational level rather than on individual parks, and provides a system for organizational learning and improvement.

Overall, none of the performance management systems has involved politicians directly in the definition of performance targets or measures, and there is a genuine lack of national political interest within the area. Standardization addresses political requirements for outsourcing and market tests of organizational performance, for example, but politicians were not directly involved in the definition of performance measures. The Green Space Award also tries to assume a political level, through increased public awareness, and acknowledgment, of the provision of a recognized and high quality service (which ought to appeal to local representatives at least). Generally speaking, there is a low level of political attention and hence institutional tension and goal conflict within the field. This is reflected in the limited involvement and interests at a political level in the definition of performance measures within park services. It is also worth noting that the development of performance management systems and practices in park services is characterized by the successive introduction of complementary systems addressing different managerial priorities. The standardization system has established a robust methodology for managing performance as defined by quantitative measures for maintenance tasks. Its limitations regarding the measurement of functional and recreational values, as well as measuring service provision at a municipal level, are addressed by systems introduced below. While the former system relies on very tight feedback mechanisms, the latter relies on loose couplings that are related more to organizational learning and development than to narrow conceptualizations and the documentation of technical efficiency.

4.3 Cross-Case Analysis

Our analysis confirms that performance management systems have been implemented increasingly in both park services and elderly care provision. In both areas, the initiation of performance management is associated with New Public Management (NPM) and a need to improve efficiency, focusing on organizational performance. In both areas, the development is propelled by the introduction of benchmarking and marketization, which forces the respective service sectors to develop quality indicators and assess performance. In this sense, similarities exist across the two sectors regarding the role of marketization in the development of key performance systems. In elderly care, the Free Choice and Common Language reforms shaped the implementation of marketization, whereas the standardization system shaped the implementation in park services. For park services, however, the challenge to define performance systems congruent with marketization trends has been addressed primarily by management, any political influence being only tenuous and related to general policies and trends promoting outsourcing in the public sector. By contrast, in the case of elderly care, politicians and government agencies have played a major role in the initiation and assessment of performance management.

Our analysis highlights a much tighter coupling with a national policy level in the definition of performance systems in elderly care compared to park services, as well as for monitoring systems in the various performance systems. In park management, we see no examples of performance systems initiated at a national level (Table 2), whereas a number of government reforms have been found to play a crucial role in shaping performance measures in elderly care (Table 1). For instance, the so-called Elderly Care Documentation Project is a comprehensive performance indicator program supported by government agencies. The involvement of politicians and public interest in elderly care seems to increase the need for control. The development of performance systems for park services has been driven more by managerial and professional needs than political involvement and ambitions. For instance, standardization systems have been developed by professionals without interest and/or direct political involvement in defining performance criteria and measures. Interestingly, later developments (e.g. the Yardstick, Parkcheck benchmark system) recognize the lack of political interest as problematic and seek to increase political involvement by providing an account which, in principle, should be able to raise the political profile. In elderly care, by contrast, the introduction of performance management is already to a large extent driven by political tension and a (political) need to signal accountability and control.

5.0 Discussion and concluding remarks

In this research we have examined the application of performance management systems across two very different public service areas (elderly care provision and park management). Given that park upkeep is a simpler service, we expected a more rigorous application of performance systems here than in the more complex field of elderly care service provision. However, we found that that the systems used in both areas converged in their characteristics, reflecting a 'one size fits

all' mentality in the use of performance management systems. The study adds to existing performance management knowledge, especially around the development of performance systems (Moynihan, 2008; Moynihan et al., 2011; Radin, 2006), by offering a description of how such systems are initiated, and the nature of the measures used in different service areas. While Radin (2006) has suggested that a one-size-fits-all, government-wide policy may lead to convergence, we argue that the convergence of performance management systems in elderly care and park services has been driven by other mechanisms. The final section of this paper outlines three propositions that could help explain the converging characteristics of performance management systems in diverse applications, and discusses how the findings contribute to related literature in this field.

Proposition 1: Policy-specific reforms contribute to the convergence of the performance management systems in the two studied public sector services. Our cross-case analysis has revealed the impact of specific, but also similar, policy reforms in each of the two areas on the development of quite comprehensive performance indicator programs. In particular, benchmark initiatives (Elderly Care Documentation Project and Yardstick, Parkcheck in park services) and privatization/marketization initiatives (Free Choice in elderly care and the Standardization System in park services) have pushed for convergence, since all initiatives require a certain level of information and performance specification. Thus, in line with existing research (Chenhall, 2007), we have found that competition calls for formalized and tight performance-control systems.

Proposition 2: Different forms of institutional pressure contribute to the convergence of performance management systems across the two studied public sector services. Another related reason for convergence in the application of performance management systems in the two cases concerned could be institutional pressure. According to DiMaggio and Powell (1983), similar management concepts are often spread across different areas and applied in similar ways through normative, mimetic and coercive mechanisms. Digging deeper into the application of management concepts, however, most studies find differences across sectors and services (Pollitt et al., 2006; Røvik, 2007). Similarly, in our study of performance management, we found a number of differences in the two sectors being scrutinized (e.g., politicians play a major role in the initiation of performance measurement in elderly care, whereas they are almost entirely absent in the equivalent process in park services), but the broad conclusion from our cross-case analysis suggests tendencies towards convergence in the characteristics of performance management systems across the two sectors. In elderly care, the pressure towards a comprehensive performance management system includes aspects which are reminiscent of a coercive mechanism, as a variety of government-initiated reforms have played a vital role in shaping performance measures. In park services, the pressure recalls more normative mechanisms, as managerial and professional needs have driven performance measures. The outcome has proved quite similar however, as extensive performance systems have been adopted in each area.

Proposition 3: While it may be more obvious to introduce a rigorous performance system with respect to a simpler service (e.g., park services) than more complex services (such as elderly care), this may be outweighed, as elderly care is characterized by high political attention in which performance systems may be a feasible tool for signaling accountability and control. In line with existing research (Askim 2007, 2009; Askim & Hanssen, 2008), we found significant differences in the political attention given to performance management system in our two sectors (high in elderly care, low in park services). Thus, strong political attention may entail tighter measures, and performance management systems may attract substantial attention if the area is highly politicized. This proposition adds to discussions around political interest in performance management systems. It implies that political attention may be just as important as task complexity when trying to understand the similarities and differences between characteristics of performance management systems; and emphasizes the extent to which performance management systems require political input to be managed (Pollitt, 2013).

In closing, we will discuss whether the identified performance management systems in park services and elderly care represent an optimal 'fit' in their specific context. As seen in our analysis of the public sector, organizations (some more than others!) are characterized by the need to comply with centralized policies and controls. In this case, organizations may adopt structures and processes that are less efficient, as seen from a selection/fit explanation in a competitive environment. In the case of elderly care provision, the development of performance management may have resulted in the adoption of less than optimal PM systems (reflected in the critique), which have subsequently been disseminated within this sector. In the case of park services, this has allowed for a more decentralized approach to the development of performance management systems than in elderly care. However, the problem then becomes the adoption of efficient performance management systems *across* the sector – models may only be adopted by a few organizations, as there are no competitive pressures for local adoption – or strong institutional mechanisms (= centralized coordination) for dissemination. Both situations (weak and strong levels of political attention) have their pitfalls. In complex settings, such as elderly care, with ambiguous goals and a politicized environment, the number of measures may easily increase in order to ensure accountability to various stakeholders (Moynihan, 2008; Pollitt, 2003). Thus, the measures are in danger of becoming just as complex and contradictory as the interests and preconceptions of different actors, therefore providing no directions for performance.

While the results bring us somewhat closer to an understanding of performance management, the analysis also raises a number of questions requiring further exploration. First, the categorization of the sectors under analysis (simple vs. complex services) may be too crude. For example, there are also complex tasks in technical services (e.g. public values associated with green spaces and their use), just as there are simple tasks associated with elderly care (e.g. cleaning). In this sense, cleaning and grass cutting are comparable tasks, so similar performance measures could be applied. Second, we are not able to conclude which of our propositions

holds most water. This would require a more theoretical, testable design, in which cases are selected systematically to test the theoretical perspectives. In order to further understand and explain similarities and differences in performance management, future research will benefit from more comparative studies and the cross-comparison of more public services.

REFERENCES

Andersen, B.R. 1999. Ældrepolitik på *afveje* [Elder policy on the wrong track]. Copenhagen: Fremad.

Askim, J. 2007. How do politicians use performance information? An analysis of the Norwegian local government experience, *International Review of Administrative Science*, 73(3): 453-472.

Askim, J. 2009. The demand side of performance measurement: Explaining councilors' utilization of performance information in policymaking, *International Public Management Journal*, 12(1): 24-47.

Askim, J. & Hanssen, G.S. 2008. Councilors' receipt and use of citizen inputs: Experience from Norwegian local government, *Public Administration*, 86(2): 387-409.

Aucoin, P. 1990. Comments assessing managerial reform, *Governance*, 3(2): 197-204.

Bell, S., Hamilton, V., Montarzino, A., Rothnie, H., Travlou, P. & Alves, S. 2008. *Greenspace and quality of life: A critical literature review.* Greenspace Scotland: Stirling.

Bevan, G. & Hood, C. 2006. What is measured is what matters: Targets and gaming in the English public care system, *Public Administration*, 84(3): 517–538.

Behn R. 2003. Why measure performance? Different purposes require different measures, *Public Administration Review*, 63(5): 586–606.

Bjørnholt, B. 2006. *Hvorfor kvalitetsundersøgelser – en analyse af kvalitetsundersøgelsers initiering, anvendelse og virkninger*. Ph.D. dissertation, Department of Political Science, Aalborg: Aalborg University.

Bovens, M., T. Schillemans & 't Hart, P. 2008. Does public accountability work? An assessment tool, *Public Administration,* 86(1): 225-242.

Boyne, G.A., Farrell, C., Law, J., Powell, M. & Walker, R.M. 2003. *Evaluating public management reforms*. Buckingham: Open University Press.

Chenhall, R.H. (2007). Theorizing Contingencies in Management Control Systems Research. In C.S. Chapman, A.G. Hopwood & M.D. Shields (eds.), *Handbook of Management Accounting Research* (pp. 163-205). Oxford: Elsevier.

Curristine, T., Lonti, Z. & Joumard, I. 2007. Improving public sector efficiency: Challenges and opportunities, *OECD Journal on Budgeting,* 7(1): 1-41.

De Bruijn, H. 2002. *Managing performance in the public sector*. London: Routledge.

DiMaggio, P.J. & Powell, W.W. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields, *American Sociological Review,* 48(2): 147-160.

FOA 2008. *Registrering og dokumentation i hjemmeplejen – indtryk fra 7 kommuner* [Documentation in home care – evidence form seven municipalities]. Udarbejdet af Bureau 2000 for FOA.

Green Space Award. 2013. *The Nordic Green Space Award*. Available at http://www.greenspaceaward.com (Accessed 25 January 2014).

Gustavsson, R., Hermy, M., Konijnendijk, C.C. & Steidle-Schwahn, A. 2005. Management of urban woodland and parks: Searching for creative and sustainable concepts. In C.C. Konijnendijk, K. Nilsson, T.B. Randrup & J. Schipperijn (Eds.), *Urban forests and trees*. Berlin: Springer.

Hansen, M.B. 2011. Antecedents of organizational innovation: The diffusion of New Public Management into Danish local government, *Public Administration*, 89: 285-306.

Hansen, M.B. & Vedung, E. 2010. Theory-based stakeholder evaluation, *American Journal of Evaluation*, 31(3): 295-313.

Halachimi, A. 2005. Performance measurements is only one way of managing performance, *International Journal of Productivity and Performance Management*, 54(7): 502-516.

Hood, C. 1991. A public management for all seasons, *Public Administration,* 69(1): 319.

Hood, C. 2006. Gaming in the target world: The target approach to managing British public services, *Public Administration Review*, 66(4): 515-521.

Jansson, M. & Lindgren, T. 2012. A review of the concept 'management' in relation to urban landscapes and green spaces: Toward a holistic understanding, *Urban Forestry & Urban Greening,* 11(2): 139-145.

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,* 25(1): 9-17.

Jones, R. 2000. Managing the green spaces: Problems of maintaining quality in a local government service department, *Managing Service Quality,* 10(1): 19-31.

Juul, J.O., Bjerregaard, S., & Dam, T. 1998. *Kvalitetsbeskrivelse for drift af grønne områder.* Frederiksberg: Skov & Landskab, KVL.

Lindholst, A.C. 2009. Contracting-out in urban green-space management: Instruments, approaches and arrangements, *Urban Forestry & Urban Greening,* 8(4): 257-268.

Lindholst, A.C. & Bogetoft, P. 2011. Managerial challenges in public service contracting: Lessons in green-space management, *Public Administration*, 89(3): 1036–1062.

Lindholst, A.C., Jensen, M. & Konijnendijk, C.C. 2012. Green Space Award: Et nyt kvalitetsstempel for parker og grønne områder, *Teknik & Miljø: Stads- og Havneingeniøren,* 2012(8): 32-33.

Lindholst, A.C., Dempsey, N. & Kreutz, S. 2014. The politics of place-keeping. In N. Dempsey, H. Smith & M. Burton (Eds.), *Place-keeping: Open space management in practice (pp. 30-51)*. London: Routledge.

Lindholst, A.C., Konijnendijk, C.C., Sullivan, S. & Fors, H. 2015. The Inherent Politics of Managing the Quality of Urban Green Spaces. *Planning Practice and Research* (forthcoming).

Locke, Edwin A. & Latham, Gary P. 2002. Building a practically useful theory of goal setting and task motivation, *American Psychologist*, 57(9): 705-717.

McLean, I., Haubrich, D. & Gutiérrez-Romero, R. 2007. The perils and pitfalls of performance measurement: The CPA regime for local authorities in England, *Public Money & Management*, 27: 111-117.

Ministry of Social Affairs. 2007. *Dokumentationsprojekt på ældreområdet*. Socialministeriet, Denmark.

Moynihan, D.P. 2008. *The dynamics of performance management: Constructing information and reform.* Washington, DC: Georgetown University Press.

Moynihan, D.P., Fernandez, S., Kim S., LeRoux, K.M., Piotrowski, S.J., Wright, B.E. & Yang K. 2011. Performance regimes amid governance complexity, *Journal of Public Administration Research*, 2: 141-155.

Nielsen, J.A. & Andersen, J.G. 2006. *Hjemmehjælp. Mellem myter og virkelighed*. Odense: University Press of Southern Denmark.

Nuppenau, C. 2009. *Marketizing municipal park management organisations in Denmark: A study of policy-making and organisational change in the period 1985-2005*. Frederiksberg: University of Copenhagen.

Pollitt, C. 2006. Performance management in practice: A comparative study of executive agencies, *Journal of Public Administration Research and Theory*, 16(1): 25-44.

Pollitt, C., Van Thiel, S. & Homburg, V. (eds) 2006. *The New Public Management in Europe: Adaptations and alternatives*. Basingstoke: Palgrave/Macmillan.

Pollitt, C. & Bouckaert, G. 2004. *Public management reform: A comparative analysis.* New York: Oxford University Press.

Pollitt, Christopher (2013). The logics of performance management. *Evaluation*, 19: 346-363

Radin, B. 2006. *Challenging the performance movement: Accountability, complexity and democratic values*. Washington, DC: Georgetown University Press.

Radnor, Z.J. & Barnes, D. 2007. Historical analysis of performance measurements and management in operation management, *International Journal of Productivity and Performance Management*, 56(5/6): 384-396.

Randrup, T.B. & Persson, B. 2009. Public green spaces in the Nordic countries: Development of a new strategic management regime, *Urban Forestry & Urban Greening,* 8(1): 31-40.

Rostgaard, T. 2006. Constructing the care consumer: Free choice of home care for the elderly in Denmark, *European Societies*, 8(3): 443-463.

Rostgaard, T. 2012. Quality reforms in Danish home care: Balancing between standardisation and individualisation, *Health & Social Care in the Community*, 20(3): 247-254.

Røvik, K.A. 2007. *Trender og translasjoner. Ideer som former det 21.* Århundrets *organisasjon,* Universitetsforlaget, Oslo.

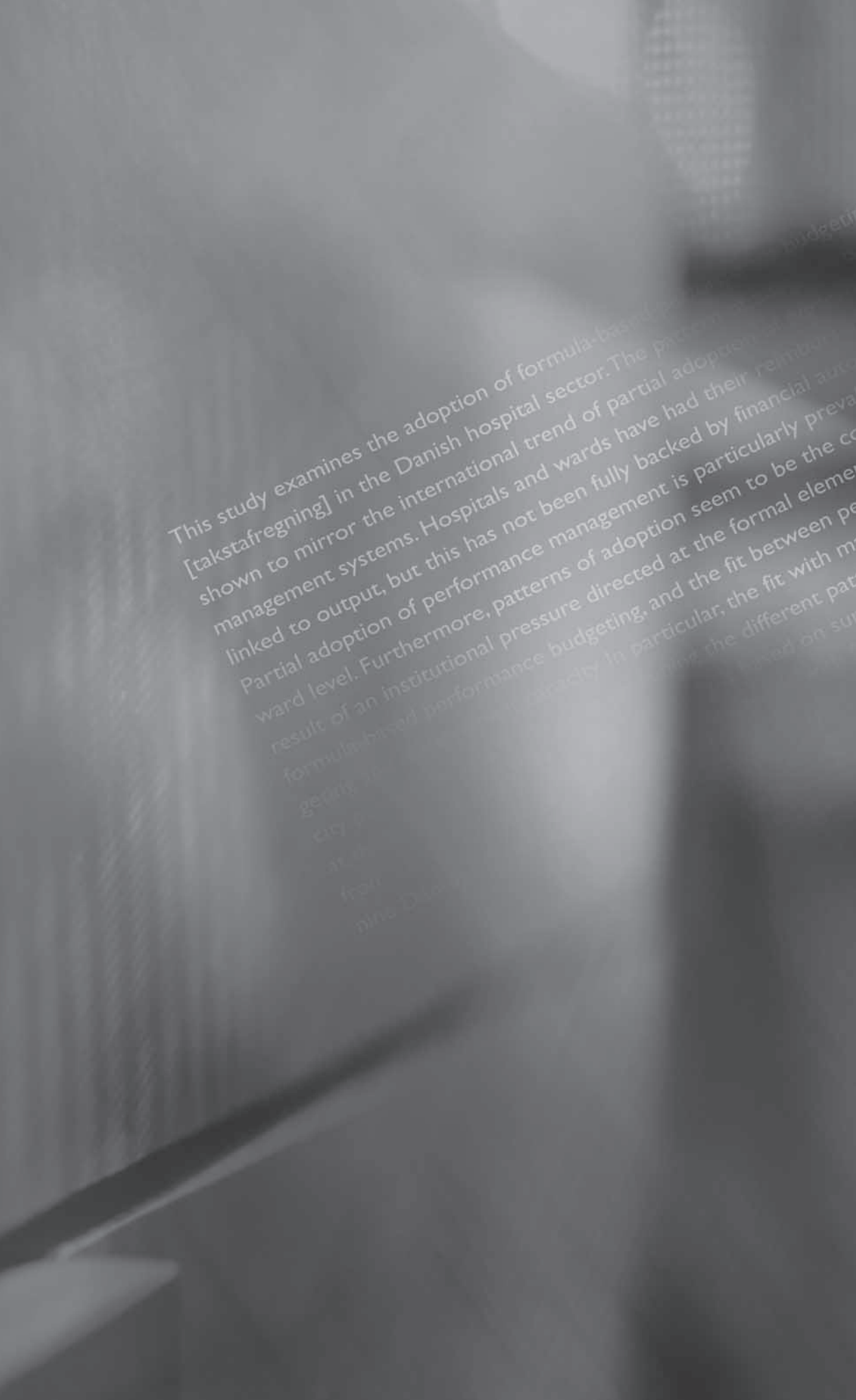
Taylor, J. 2009. Strengthening the link between performance measurement and decision making, *Public Administration,* 87(4): 853-887.

Van Dooren, W., Bouckaert, G. & Halligan, J. 2010. *Performance management in the public sector*. Routledge: Great Britain.

Vedung, E. 2010. Four waves of evaluation diffusion, *Evaluation*, 16: 263.

Yardstick. 2012. Parkcheck general prospectus 2012. Available at www.yardstickglobal.org (accessed January 25, 2014).

Yin, R.K. 2008. *Case study research: Design and methods*, Fourth Edition*.* Thousand Oaks: SAGE.


This study examines the adoption of formula-
[takstafregning] in the Danish hospital sector. The
shown to mirror the international trend of partial
management systems. Hospitals and wards have had their
linked to output, but this has not been fully backed by financial
Partial adoption of performance management is particularly
ward level. Furthermore, patterns of adoption seem to be the
result of an institutional pressure directed at the formal
budgeting, and the fit between
In particular, the fit with

The partial adoption of performance budgeting at Danish hospitals

Mads Leth Felsager Jakobsen

Abstract

This study examines the adoption of formula-based performance budgeting *[taksstafregning]* in the Danish hospital sector. The pattern of adoption is shown to mirror the international trend of partial adoption of performance management systems. Hospitals and wards have had their reimbursements linked to output, but this has not been fully backed by financial autonomy. Partial adoption of performance management is particularly prevalent at the ward level. Furthermore, patterns of adoption seem to be the combined result of an institutional pressure directed at the formal elements of formula-based performance budgeting, and the fit between performance budgeting and management capacity. In particular, the fit with management capacity provides a strong basis for explaining the different patterns of adoption at the hospital and ward levels. This study is based on survey data gained from 26 Danish hospitals in 2007, and interviews with financial officials at nine Danish hospitals in 2011.

1. Introduction

Use of performance measurement and management systems for improving organizational performance has dominated discussions about public sector management in recent decades (Pollitt and Bouckaert, 2011; Robinson, 2002). However, formal adoption of such systems does not automatically equate to real change. Not only have performance management systems often failed to deliver the expected benefits (Moynihan, 2008; Meier and O'Toole, 2009; Hood and Peters, 2004), but also in many cases they have not been fully adopted. In particular, it has been difficult to secure the financial and operational autonomy envisaged in the doctrine of performance management (Utz, 2010; De Araújo and Branco, 2009; Binderkrantz and Christensen, 2009; Moynihan, 2008; Goldfinch, 2006, Pettersen and Nyland, 2006; Pettersen 1999). The result is a pattern of »partial adoption« or »half-hearted managerialism« (Moynihan, 2008; Hood, 2002).

This paper examines the extent to which a specific form of performance management – formula-based performance budgeting – has been adopted at Danish hospitals and on hospital wards. Formula-based performance budgeting is a very

strong form of performance management, as it attaches financial incentives to performance by linking the performance of organizations to their funding (Bech, 2004; Robinson, 2002: 18-20; Jegers et al., 2002: 260). It has been introduced in the Danish hospital system under the heading »Activity-Based Reimbursement« *[takstafregning].*

The paper seeks to answer three questions:

1. Has the adoption of formula-based performance budgeting at Danish hospitals mirrored the international trend of partial adoption?
2. Has the adoption of formula-based performance budgeting differed between organizational levels within hospitals?
3. What are the likely explanations for the observed patterns of adoption?

By answering these questions, the paper makes three contributions to research in this field. First, it provides a first broad systematic account of how formula-based performance budgeting has been adopted at Danish hospitals, at both the hospital and ward levels. Such an account is especially important in relation to wards, since these are central to the performance of hospitals and yet our knowledge about this level is limited. The main empirical finding is that partial adoption is prevalent at both levels; either management does not formally grant financial autonomy, or it does not (as far as is evident from its actions) respect this autonomy. Furthermore, there is a higher level of partial adoption (and outright rejection) of formula-based performance budgeting at the ward level than at the hospital level.

Second, this empirical analysis strengthens the finding from the performance management literature that partial adoption is a very common phenomenon. It is a non-trivial contribution to the literature too, because, in contrast to much of the other literature on performance management (see Moynihan, 2008), it focuses on a strong performance management system based on economic incentives, and examines partial adoption in terms of not just formal but also behavioral adoption.

Third, by combining quantitative comparisons of adoption at hospital and ward levels with interview data on reasons for adopting formula-based performance budgeting for wards, the paper finds strong indications that the different adoption patterns at the hospital and ward levels are due to variations in the fit between formula-based performance budgeting and management capacity. The strong incentives of formula-based performance budgeting is more difficult to align with achieving organizational goals at a ward compared to a hospital level. This explanation is somewhat at odds with traditional explanations of partial adoption which emphasize how institutional pressure attaches symbolic value to only some elements of performance management systems (eg performance) but not to other parts (such as autonomy) (Moynihan, 2005).

The analysis is based on a 2007 survey of 26 Danish hospitals about their adoption of formula-based performance budgeting. This is combined with the findings

from qualitative interviews with financial officials from nine hospitals in 2011 on the use of formula-based performance budgeting for wards.

The paper first develops the analytical framework and derives empirical expectations, before going on to outline the data for analysis. It then provides empirical analysis of the patterns of adoption and likely explanations, from which we draw our conclusions and propose suggestions for subsequent research.

2. Analytical framework and empirical expectations

Broadly defined, performance management is »...a system that generates performance information through strategic planning and performance measurement routines and that connects this information to decision venues, where, ideally, the information influences a range of possible decisions« (Moynihan, 2008: 5). It includes management systems that have different forms – for example whether or not performance results are made public, and whether or not they are linked to incentives (Moynihan, 2008; Propper & Wilson, 2003: 253-254). The doctrine of performance management states that results will improve when public organizations are given clear goals, incentives to pursue the goal, and autonomy to find their own best way to accomplish this (Moynihan, 2008; Hood, 2002).

This paper looks at a specific type of performance management: formula-based performance budgeting. That is a management system based on: (1) a formula linking performance results to reimbursement, and (2) financial autonomy with respect to means and economic dispositions (Robinson, 2007a: 54; Davies, 1999: 151; Jegers et al., 2002: 60). If an organizational unit submitted to formula-based performance budgeting delivers a poorer performance than expected, it will receive less money than expected – and the opposite if it delivers a performance that is better than expected. Furthermore, the unit has autonomy to use its resources as it sees fit; if it delivers a surplus, it can keep this and use it within the unit. However, if for instance a deficit should occur due to bad management, the unit must also recover this deficit itself in the subsequent period. Performance is thus financially incentivized.

In recent decades, such formula-based performance budgeting systems have been widely adopted in hospital sectors around the world (Gilardi et al., 2009). Performance is measured as activity within the diagnosis-related group (DRG) system that groups patients with medically-related and economically-comparable diagnoses (Cots et al., 2011: 75; Ankjær-Jensen et al., 2006).

2.1 Partial adoption of formula-based performance budgeting

When an organization is introduced to a management concept such as formula-based performance budgeting, it can respond in one of three ways. It can reject the concept altogether, it can fully adopt the concept and turn it into practice, or it can do something in between. Partial adoption describes these »in-between« responses. A more explicit conceptualization of partial adoption can be derived from institutional theory (Oliver, 1991) which distinguishes between formal and

behavioral adoption. This is shown in Table 1 where four types of responses are outlined.

'Acquiescence' means that the organization both formally adopts the new concept and changes its behavior accordingly (Oliver, 1991: 152-153). Hence, the organization sticks to the doctrine behind the concept. With formula-based performance budgeting that would imply not only the formal adoption of formula-based reimbursement and financial autonomy, but also that these were put into practice.

Table 1. Inward organizational responses to a new management concept

		Formal adoption	
		Yes	No
Behavioral adoption	Yes	Acquiescence	Bottom-up
	No	Avoidance	Defiance

In the case of 'Avoidance', the organization separates the formal adoption of the management concept from actual behavioral adoption, which is absent. This is one form of partial adoption – ie where the new concept is formally adopted but not put into practice. 'Defiance' means that the management concept is neither formally nor behaviorally adopted. Between avoidance and defiance there is another case of partial adoption, however, which is not shown in Table 1. Here, the organization formally adopts some, but not all of the elements of performance budgeting – for example formula-based reimbursement but not financial autonomy. This can be called 'Partial Defiance'. Finally, an organizational response characterized by behavioral adoption but no formal adoption would indicate a bottom-up response from the lower echelons of the organization.

How organizations such as Danish hospitals respond to the concept of formula-based performance budgeting can be conceptualized as a question about whether they adopt three different aspects of the concept (see Robinson, 2007a; Bech et al., 2006; Serritzlew, 2006; Biørn et al., 2003):

1. A formula linking output and reimbursement;
2. Formal transfer clauses for surpluses and deficits;
3. No behavioral exemptions from such formal transfer clauses.

The first two aspects measure the formal linkage between performance and finances and the formal establishment of financial autonomy, while the last measures whether financial autonomy is put into practice.

The incentives created by the system are the result of all three aspects. The formula establishes the link between performance and reimbursement. As is typical of the formulae in the Danish hospital sector, it sets a fixed rate of reimbursement that a unit such as a hospital or ward receives for each increment of performance. The scale of this rate indicates the strength of the formula. The more substantial the rate for each extra increment of output, the stronger the incentives faced by

the hospital or ward to deliver performance. If, however, the units do not enjoy the fruits of a surplus or incur the trouble of a deficit, the formula does not present adequate financial incentives to drive performance. If deficits are compensated, the measures also have less of a consequence. That is why not only formal adoption, but also the absence of exemptions, are important for the efficacy of formula-based performance budgeting.

Table 2 shows how different responses to the concept of performance budgeting can be categorized.

Table 2. Categorizing organizational responses to formula-based performance budgeting

		Acquiescence	Partial adoption		Defiance
			Avoidance	Partial defiance	
Formal adoption	Linkage (formula) between rewards and output	+	+	+	÷
	Formal transfer clauses	+	+	÷	÷
Behavioral implementation	No exemptions from transfer clauses	+	÷	÷	÷

Acquiescence requires not only the formal adoption of reimbursement of output and transfer clauses, but also the refusal to permit exemptions from these transfer clauses. Partial adoption happens either through avoidance (where output reimbursement and transfer clauses are formally adopted but not put into practice) – or partial defiance (where transfer clauses are not formally adopted nor put into practice). If a formal link has not even been established between output and reimbursement, we have a case of defiance.

2.2 Explaining partial adoption and its variations

Two broad arguments are highlighted in the performance management literature to explain the adoption of performance management: institutional pressure and management capacity (Moynihan, 2008). The 'institutional pressure' perspective focuses on whether external factors coerce and entice the organization to adopt performance management, while the 'management capacity' perspective focuses on whether performance management helps public managers realize organizational goals.

2.2.1 Institutional pressure

Institutional pressure on an organization arises from the institutional environment of the organizations. This environment consists of actors interacting within the same societal field such as a policy sector. Such fields can give rise to strong symbolic pressures from fashions and norms about what constitutes the right

management systems (Zucker, 1987: 449-450). Without at least formal adoption of 'fashionable' management techniques, organizations become vulnerable to external criticism (Brunsson, 2002; Meyer & Rowan, 1977). Moynihan (2005) has found that the symbolic value of performance management systems has been a key factor behind the adoption of performance management in the US.

Institutional pressures can, however, also be coercive – for instance within hierarchical relationships between political authorities and subordinate organizations (Oliver, 1991: 167-170; DiMaggio & Powell, 1983). Lega and Vendramini (2008) have shown how the adoption of performance management systems in the Italian hospital system was driven both by symbolic and top-down coercive pressures. The stronger and more homogenous the symbolic and coercive pressures are, the more organizations can be expected to adopt the management concept (Scott & Davis, 2007: 131ff; DiMaggio & Powell, 1983).

2.3 Institutional pressure in Danish hospitals

The Danish hospital sector is typical of the Scandinavian welfare state – in that it is tax-funded, has universal access, is decentralized, and is publicly owned. Until 2006, public hospitals were owned and managed by 14 democratically-elected county councils, which in 2007 were replaced by five democratically-elected regional councils. Above this level, the central government sets the overall legislative framework for the sector, but it is not directly involved in the management of hospitals (Vrangbæk, 2012; Strandberg-Larsen et al., 2007: 48). So, the county and regional councils have had wide discretion in their management of hospitals (Ankjær-Jensen et al., 2006: 1; Damgaard & Pallesen, 1998: 133-134). At the hospitals, the management has hierarchical authority to run wards and other hospital functions within the limits set by the region (Christensen et al., 2011: 176-177). Applying the 'institutional pressures' perspective to the Danish public hospitals sector gives rise to specific expectations about the adoption of formula-based performance budgeting.

From the turn of the century, Danish county (and regional) councils and hospitals have been exposed to a strong institutional pressure to adopt formula-based performance budgeting (Ankjær-Jensen et al., 2006: 2). This has been justified by its claimed positive effect on hospital activity and efficiency (Indenrigs- og Sundhedsministeriet, 2005: 22; Ministry of Finance, 2003: 15).

The pressure has been coercive in relation to some aspects of formula-based performance budgeting. Through legislation and annual budget agreements, the central government has demanded the formal use of formulae linking output and hospital reimbursement. These formulae were to cover 10 percent of hospital reimbursement in 1999, 20 percent in 2004, and 50 percent in 2007 (Jakobsen, 2010). There has been no coercive pressure to adopt financial autonomy in the form of transfer clauses; it has been recommended in the annual agreements between the counties/regions and the central government, however, that hospitals should have increased levels of autonomy (Ministry of Finance, 2003: 38). In a few coun-

ties and regions, the councils have also coerced hospitals to adopt some kind of formula-based reimbursement for their wards (Dørken et al., 2012: 74-75; H:S et al., 2005: 61). The coercive pressures for formula-based performance budgeting is shown in the right column of Table 3.

There are also strong indications that the coercive pressure has been supplemented by a symbolic pressure to formally adopt formulae linking output and reimbursement and financial autonomy at both a hospital and ward level. A range of official reports has recommended formula-based performance budgeting as the future management model for the Danish hospital sector (Møller Pedersen et al., 2006: 37-39). During the same period, formula-based performance budgeting became widely used in most western countries, indicating that it had become the fashionable way to manage hospitals (Wiley, 2011: 3; Jakobsen, 2010; Gilardi et al., 2009). This pressure is illustrated in the centre (and right) column of Table 3.

Symbolic pressures can only encompass what is visible to the actors in the institutional environment, however. This is not the case for exemptions to transfer clauses within the Danish hospital sector. Exemptions are not part of formal descriptions of performance budgeting systems. Both at the counties/regions and at the hospitals, management has discretionary power to allow for exemptions to transfer clauses at hospital and ward levels respectively, without embarking on a publicly visible and salient process of formally changing the rules. If the exemptions had required legislative changes, as is the case with private schools in Denmark for instance, this would have been much more troublesome and visible to the general public and thus a potential subject for institutional pressure (see Serritzlew, 2006). This lack of pressure in relation to exemptions is shown in the left column of Table 3.

Table 3. Institutional pressure and fit with management capacity

		Institutional pressure		
		No explicit pressure	*Symbolic*	*Coercive (and symbolic)*
Fit with management capacity	*Higher*	Exemptions from transfer clauses (hospitals)	Formal transfer clauses (hospitals)	Reimbursement of output (hospitals)
	Lower	Exemptions from transfer clauses (wards)	Formal transfer clauses (wards) Reimbursement of output (most wards)	Reimbursement of output (a few wards)

Horizontally, Table 3 illustrates how the institutional pressure varies between formal formula-based reimbursement, formal transfer clauses and exemptions from these transfer clauses. We have:

- coercive and symbolic pressure for adoption of a formula linking performance to reimbursement;
- symbolic pressure for formally adopting transfer clauses; and
- no pressure for not exempting from transfer clauses.

Based on the 'institutional pressures' perspective, we should expect partial adoption of formula-based performance budgeting in the Danish hospital sector, because the pressure has not encompassed exemptions from transfer clauses. Furthermore, we should expect slightly increased defiance at wards compared to hospitals, as only few wards have been subjected to a coercive pressure to adopt a formula linking output and reimbursement.

2.3.1 Management capacity

Assuming that public organizations attempt to fulfill their tasks (Scott & Davis, 2007: 125), the management of an organization is more likely to adopt or resist new management concepts the more it perceives them as either enhancing or reducing their capacity to do this (Oliver, 1991: 165; Ouchi, 1980). Management capacity is simply the ability of the management of a given organization to make the organization fulfill its tasks (definition inspired by Donahue et al., 2000: 384). In this way, this perspective emphasizes the instrumental value of formula-based performance budgeting.

Proponents of formula-based performance budgeting, and performance management more generally, argue that these management systems enhance management capacity (Robinson, 2007a: 82; Moynihan, 2008: 36). They do so by aligning the behavior of the organization with managers' objectives. For instance, by changing the formula so that reimbursement is linked to specific forms of performance, the management can change which kinds of behavior are attractive and non-attractive to perform for the units within the organization. In the hospital sector, formula-based performance budgeting was designed to make hospitals and wards more cost-efficient and productive – thus an obvious response to problems of inefficiency and long waiting times (Robinson, 2007a; Jakobsen, 2010).

There are also potential downsides of formula-based budgeting to management capacity, however. It is for this reason that the 'management capacity' perspective does not necessarily imply that acquiescence is the expected standard response to formula-based performance budgeting. Three distinct characteristics of public-sector organizations such as hospitals can make formula-based performance management systems challenge management capacity. First, the organizations concerned often have multiple goals – including quality, efficiency, cost control, and activity, for example. Formula-based performance budgeting can cause goal displacement if only some dimensions (such as 'activity') are incentivized but others (such as quality) are not (Or & Häkkinen, 2011: 117; Robinson, 2007b).

Second, the organizations have interdependent production processes with many actors involved in the production of services. Reimbursement of individual units

involved in interdependent production (Street et al., 2011: 111) creates incentives for the units to focus narrowly on the specific part of the production process for which they are reimbursed, and not coordinate with other relevant production units. This hampers overall goal achievement.

Third, the organizations have large output heterogeneity where the costs of similar cases (for example similar diagnoses) vary significantly. Variations in the actual cost of cases with similar rates of reimbursement create incentives for selection – for example creaming and dumping – which are at odds with goals such as equity and efficiency (Cots et al., 2011: 85; Andersen & Jakobsen, 2011; Smith, 2007).

Such dysfunctional effects of formula-based performance budgeting have also been empirically documented in the hospital sectors in many other countries (O'Reilly et al., 2012; Bevan & Hood, 2006; Busse et al., 2006).

2.4 Management capacity and Danish hospitals

The more an organization is able to reap the benefits of formula-based performance budgeting while mitigating its downfalls, the better the fit between management capacity and formula-based performance budgeting. This is shown vertically in Table 3 for Danish hospitals and wards.

The structure of hospitals makes them less vulnerable than wards to the downsides of formula-based performance budgeting. Most patients are treated within one rather than several hospitals (for data from the Danish hospital sector, see Draborg et al., 2009: 17); only for patients requiring more specialized treatment is it typically necessary to involve other hospitals (Damgaard & Pallesen, 1998: 132). When reimbursement for output happens at a hospital level, it is therefore only for the few patients treated at multiple hospitals that problems related to interdependent production arise. Within hospitals however, most patients are treated through interdependent production processes involving several clinical wards and auxiliary functions – for example laboratories and X-ray departments (Bendix et al., 2012: 69-70; Christiansen, 1996: 117-118). Reimbursing output at the ward level, then, means that the individual units involved in interdependent production are reimbursed separately and not collectively. This creates problems of coordination that are less prominent at a hospital level.

The size of hospitals also makes them less vulnerable to output heterogeneity. The more patients that are treated at a hospital or a ward, the better the unit can mitigate output heterogeneity (Cots et al, 2011: 86). What the unit loses on the roundabouts can be gained on the swings, if the number of cases is large enough. For single wards, cases that are not fully reimbursed are more detrimental to the economy than they are to hospitals that have many other cases to balance these expensive cases (Robinson, 2007c: 300).

As the negative implications of formula-based performance budgeting grow and the fit with management capacity diminishes, we should expect formula-based per-

formance budgeting to be adopted in a way that creates weaker incentives (Cots et al., 2011: 90; Smith, 2007). This reduces the positive effects of formula-based performance budgeting, but it also reduces its unintended negative side-effects. By lowering rates or not formally granting autonomy, the management limits the side-effects and can redistribute surpluses to mitigate problems arising elsewhere in the system. They can do the same by allowing exemptions from formal transfer clauses – for instance by making it possible to reabsorb a surplus from one unit for use to cover a deficit in another unit, where it is considered unacceptable that the unit recovers the deficit by itself.

Consequently we should expect more defiance or partial adoption at a ward level compared to a hospital level. This contrasts with the 'institutional pressure' perspective that only predicted differences between the hospital and the ward level in the formal adoption of a formula linking output and reimbursement.

3. Data

Our quantitative data on the adoption of performance management at Danish hospitals and wards was gathered via an email survey issued to the financial sections of all Danish hospitals. The relevant survey questions are presented as part of Table 4. Of the 30 hospitals approached (ie all Danish hospitals in 2007), four chose not to participate. Data also covers 27 orthopedic wards (a few hospitals had more than one ward), which are included in the analysis to represent the ward level. Due to their many elective patients, which constitute a feature suitable for formula-based performance budgeting, orthopedic wards are a likely case for acquiescence. The survey data was gathered between 15 November, 2007 and 1 July, 2008 (see data report, Jakobsen, 2009).

To assess the causes of differences between the hospital and ward levels, the paper also draws on the findings of an interview study conducted in 2011with the financial officials of nine Danish hospitals[1]. The interviews included highly relevant questions for this study such as »How are wards reimbursed?«, and »Why are they reimbursed in this way?«, with follow-up questions on management capacity and coercion. The interview data have been content-coded (see Table 6) and analyzed following the central principles of display analysis: inclusion of all contradicting statements and authenticity when statements are condensed (Miles and Huberman, 1994).

4. Patterns of adoption

Whether hospitals and wards have reimbursement that is related to output is measured in two ways.

First, the average of the reimbursement rate for output below and above the baseline, is displayed in Table 4. The average is used because some counties/regions and hospitals use different reimbursement rates for output before and after a specified level of baseline production.[2] All hospitals and wards use formulae that

set the rate as a share of the DRG rate, which is determined by the national health authorities.

Table 4. Adoption of formula-based performance budgeting at Danish hospitals and orthopedic wards (percentage), 2006

Hospitals	Average DRG reimbursement rate[a]	53	n = 24
	Formula-based reimbursement of output	100	n = 24
	Formal transfer clause[b]	72	n = 25
	No exemptions from clause (2005-2007)[c]	50	n = 20*
Orthopedic wards	Average DRG reimbursement rate[a]	35	n = 24
	Formula-based reimbursement of output	80	n = 24
	Formal transfer clause[b]	38	n = 26
	No exemptions from clause (2005-2007)[c]	33	n = 15*

a »What share of the DRG rate did the unit receive per patient before the baseline level was reached?« and »What share of the DRG rate did the unit receive per patient after the baseline level was reached?« (author's own translation).
b »How large a surplus or deficit could be transferred between budget years?« (author's own translation).
c »Have there been exempted from the transfer clauses?« (author's own translation).
* This number also includes wards that had formal transfer clauses in either 2005 or 2007, but not in 2006.

On average, hospitals received 53 percent and wards 35 percent of the DRG rate for their output, which is also a statistically significant difference ($p = 0.0022$).[3] Increasing or lowering output thus has larger financial implications at the hospital than at a ward level. However, the different levels of reimbursement may partly reflect the fact that the hospitals deliver services to the individual wards and thus have to levy a tax on the reimbursement rate before it is passed on to the wards. This is also pointed out in the interviews with hospitals' financial officers (see Table 6, below). There is no accessible data to determine the precise level of this tax.

Second, a more comparable parameter between hospital and ward levels then is whether they have been subjected at all to a formula involving reimbursement of output (see Table 4). In 2006, counties had subjected all hospitals to this form of reimbursement. At ward level, where the hospital management makes the decisions on reimbursement, reimbursement of output was used somewhat less. Twenty of 24 orthopedic wards (80 percent) had been subject to reimbursement by output. This is not a very big difference, but it is still statistically significant ($p= 0.018$).[4] Thus, looking only at the formal adoption of a formula linking reimbursement to output, we have indications of a very high degree of adoption of formula-based performance budgeting.

Table 4 also shows the formal adoption of transfer clauses. We see a stronger pattern of adoption at the hospital than at the ward level. Seventy-two percent of the hospitals were subjected to such transfer clauses compared to 38 percent of the orthopedic wards, which is also a significant difference ($p= 0.032$).[5] It is worth noting that more than 60 percent of the wards had not been formally subjected to both of the two constituent elements of formula-based performance budgeting: formula-based reimbursement of output and financial autonomy. On the other

hand, three in four hospitals were – at least, formally – still loyal to the concept of formula-based performance budgeting.

Finally, Table 4 shows whether hospitals and wards experienced exemptions from transfer clauses in the period 2005-2007. Among those with transfer clauses, 40 percent of hospitals and 20 percent of wards with formal transfer clauses were not subject to exemptions during the period, which is not a significant difference (p= 0.2072).[6]

Giving a more precise overview of the patterns of adoption at the hospital and ward levels, the overall organizational responses are presented in Table 5.

Table 5. Organizational responses (in percentage) of Danish hospitals and orthopedic wards, 2006

Response		Hospitals	Orthopedic wards
Acquiescence		36	13
Partial adoption	Avoidance	36	26
	Partial defiance	27	39
Defiance		0	22
Total		99	100
n		22	23

The Table shows that about one third of hospitals (36 percent) have experienced either acquiescence or avoidance, while 27 percent have experienced partial defiance. Thus, 63 percent of hospitals have experienced partial adoption of formula-based performance budgeting (avoidance + partial defiance). At the ward level, 65 percent have partially adopted the practice.

Still, there are important differences between the two levels. First, the stronger version of partial adoption – partial defiance – is more widespread among wards than hospitals. Second, ward-level practice also yields five cases of outright defiance while hospital-level use doesn't give rise to any. The median response of hospitals – avoidance – is also significantly different from the median response of wards, which is partial defiance (p=0,008).[7] The main picture is thus one of partial adoption, particularly at a ward level.

5. Interpreting the patterns of partial adoption

What are the likely explanations of these patterns? In line with the literature on performance management, this paper has focused and elaborated on an 'institutional pressure' perspective and a 'management capacity' perspective.

Focusing on institutional pressure led us to expect a high degree of formal adoption of formulae linking reimbursement to output. This turned out to be the case at both the hospital and ward levels. There were a few defiant responses at the ward level, but in contrast to the hospital level there were no coercive institutional

pressures. The institutional pressure did not cover exemptions to formal transfer clauses, which allowed for a very different response at the hospital and ward levels. That some form of partial adoption was the median response in both scenarios can at least partly be ascribed to the institutional pressure then.

However, the institutional pressure perspective presents some difficulties in explaining the relative low levels of formal adoption of transfer clauses and the high level of exemptions from such clauses at the ward level. This is likely to be because this perspective does not allow for the countervailing forces that might work against adoption in cases of pressure (formal adoption of transfer clauses), or what drives adoption in cases of no pressure (exemptions from transfer clauses).

The difference between the hospitals and the wards corresponds better with the 'management capacity' perspective. We should expect a lower adoption of formula-based performance budgeting at the ward level, because the fit with management capacity is lower. This was also reflected in the overall organizational response which was more partial and defiant for wards than hospitals. This indicates that management capacity has been an important factor shaping the adoption patterns.

Whether such an interpretation is empirically valid can be further examined drawing on Table 6. This presents statements from financial officers at nine hospitals about why, in 2011, their hospitals had or had not passed on the system of formula-based performance budgeting to the ward level. If both a formula based on reimbursement of output, and on transfer clauses was passed on, the organizations are categorized as adopters in the Table. Four of the hospitals had done this, while five had not.

The findings support the interpretation that management capacity has shaped the decision to adopt or not adopt formula-based performance budgeting, and the differences in these decisions between hospital and ward levels.

In relation to institutional pressure, the interview statements indicate very little coercion from the regions. It was not the regions that had decided whether the hospitals adopted formula-based funding of their wards. Furthermore, the one hospital that perceived coercion did not consider this important for its decision to adopt formula-based performance budgeting for its wards. The fact that the regions themselves were coerced by the state to use formula-based performance budgeting of the hospitals, however, does indicate that variations in the external pressure were a scope condition for the higher degree of partial adoption and defiance at the ward level.

Still, the interview statements primarily support the interpretation that the fit with management capacity was the dominant factor behind adoption of formula-based performance budgeting. It is almost exclusively considerations relating to management capacity that are used to argue for both adoption and non-adoption

Table 6. Arguments for and against the use of formula-based performance budgeting at wards

	Have adopted (four hospitals)	Have not adopted (five hospitals)
Coercion	One hospital perceives a demand from the region	No hospitals perceive a demand from the region
Positive effects	»otherwise it makes no sense that we have to (…) create incentives at the wards«. »it really has to be adjusted to clinical practice (…) otherwise it results in some strange incentives.«	[Formula-based performance budgeting results in] »a better incentive structure, probably«.
Coordination challenges	»the formula-based performance model provides (…) anesthesia with an incentive to make sure that orthopedic surgery doesn't just sit back.« »that planning responsibility lies with those (…) who know most about it«. »conflicts like that will arise (…) often the hospital management will look into it«. »central pool (…) to level some of those injustices«.	»interdisciplinary functions … There are advanced models for that sort of thing, but it should also (…) make sense.« Formula-based performance budgeting creates »tremendous focus on whether there's a change of patient (…) between wards … or if (…) our contract is not (…) quite accurately stitched up.« »a holistic way of thinking … the wards have an added activity to be paid, but the hospital has less activity.«
Other dysfunctional incentives	The wards »can be so conscious [of the incentives] that they say it'll be on Monday (..) we'll make more money, and somehow that's inappropriate.« »many micro things (…) can have a huge importance to (…) the wards.« »We would also like if some wards were more aware of registering correctly.« »a decrease in DRG value per patient contact (…) affects them financially even though it may not be quite fair.«	»some (…) changes in activity (…) the ward is not to be blamed for«. »will it cause indications [of diagnosis] to slip?« [with formula-based performance budgeting] »with the 70-20 it hasn't (…) made any sense… it would result in some major injustices.«
The issue of size	»the wards are (…) so big that it makes sense to (…) talk about average.«	»plus 10 or (…) minus 70, there's a world of difference between doing it at the ward level and at the hospital level.« »you have (no) means of control at the ward level and without a doubt, it becomes much, much harder the smaller the unit you have to run … at the hospital level, it is somewhat easier because we have some things that go up and some that go down.«

Based on statements coded in two consecutive rounds:
Round 1 codes – »the relationship between hospital and ward« + »transfer of surplus and deficits« + »reimbursement formulae«
Round 2 codes (only if coded in round 1) – »arguments related to the use of formula-based performance budgeting at wards«

at the ward level. Both adopters and non-adopters are keenly aware of the positive effects that formula-based performance budgeting can have on activity and productivity. However, the adopters consider these effects to be so important that they merit adoption (even when taking the negative effects into account).

Many of the considerations revolve more specifically around the issue of interdependencies within the hospital and how to handle them if formula-based performance budgeting is passed on to the ward level. For some non-adopters, this is a key reason for not passing the system on to the ward level. They emphasize

problems of reimbursing the auxiliary units, as well as the issues that arise when individual wards are not aligned with the overall goals and needs of the hospital. These problems are also noted by adopters, but some find them manageable or at least not detrimental enough to outweigh the benefits. The same is the case when it comes to other possible negative effects of formula-based performance budgeting. The interviewees point to the risk of gaming, and the feelings of injustice the formula create by not giving what is considered to be fair reimbursement to all wards.

It also underlines the importance of the 'management capacity' perspective that the key mechanisms behind the expected differences between the hospital and the ward levels are clearly present. Both adopters and non-adopters are aware of the importance of the size of hospitals and wards respectively for their fit with formula-based performance management. Among adopters, one hospital official emphasizes the large size of the wards at his particular hospital, while the non-adopters point to the limited size of wards compared to hospitals.

In summary, we have good reasons to believe that considerations about the impact of management capacity are important in the decision to adopt and particularly partially adopt formula-based performance management systems in the Danish hospital sector. Institutional pressures also have a role to play, but management capacity seems more important in relation to the differences at hospital and ward levels.

We found no signs of other important explanatory factors in the interview data. This could not be ruled out a priori as the literature on policy adoption and performance management has also pointed to factors such as veto points and ideology as potential explanations for adoption of formula-based performance management (Binderkrantz & Christensen, 2009; Gilardi et al., 2009). Furthermore, considering the difference between hospital and ward levels, it does not seem likely to be the result of either ideology since the political principals are the same; nor of different veto structures since there are more potential veto points in the broad county/regional councils with many political parties than in the management teams of the individual hospitals.

6. Conclusion

In the Danish hospital sector, the dominant organizational response to the concept of formula-based performance management has been partial adoption in the form of avoidance and partial defiance. The formal adoption of formula-based funding has been widespread, but there have also been many instances at both hospital and ward level where this has not been followed by either formal adoption of transfer clauses or a practice of not exempting from such clauses. The pattern of adoption thus reflects the international trend of partial adoption of performance management systems. This is a finding that calls for further research, as studies have started to empirically show how different degrees of partial adoption shape the performance of public sector organizations (Nielsen, 2013).

The dominant organizational response, however, also differs between the hospital and ward levels. There is a stronger form of partial adoption at the ward level where the median response is partial defiance, than at the hospital level where the median response is avoidance.

Finally, the analysis has shown that likely explanations for these patterns of adoption can be found within the 'institutional pressure' and 'management capacity' perspectives that are central in performance management literature. While variations in institutional pressure would predict some form of partial adoption, it can neither account for non-adoption in cases of pressure, as with formal transfer clauses, nor for the marked difference between the organizational levels. Instead, the paper has found strong indications that this difference is the result of the way the relevant actors interpret formula-based funding to fit with their capacity to achieve organizational goals.

Overall, the study has identified patterns of adoption of formula-based performance budgeting in the Danish hospital sector as well as some likely explanations for these patterns. While these results contribute to the general literature on performance management their limitations must be kept in mind. By focusing on a highly financial and incentivized form of performance management, the conclusions cannot be generalized to other less incentivized forms of performance management. While the pattern of partial adoption has been found to be rather general in the performance management literature, it is quite likely that less incentivized forms of performance management would have a better fit with management capacity at a ward level than is the case with formula-based performance budgeting. Still, this is an issue that merits further research. Furthermore, as the paper is based on cross-sectional data and utilizes simple bivariate analyses, it cannot make very strong claims on causality. But the results definitely call for further research, particularly into the impact of management capacity.

References

Andersen, L.B. & Jakobsen, M. 2011. Does ownership matter for the provision of professionalized services? Hip operations at publicly and privately owned clinics in Denmark, *Public Administration,* 89, 956-974.

Ankjær-Jensen, A., Rosling, P. & Bilde, L. 2006. Variable prospective financing in the Danish hospital sector and the development of a Danish case-mix system, *Health Care Management Science,* 9(3), 259-68.

Binderkrantz, A.S. & Christensen, J.G. 2009. Governing Danish agencies by contract: From negotiated freedom to the shadow of hierarchy, *Journal of Public Policy,* 29(1), 55-78.

Bech, M. 2004. *Choice of Hospital Reimbursement Scheme: Incentives and Tradeoffs.* Odense: University Press of Southern Denmark.

Bech, M., Lauridsen, J. & Pedersen, K. 2006. Giver øget brug af takststyring i sygehusvæsenet øget produktivitet?, *Nationaløkonomisk Tidsskrift*, 144, 326-342.

Bendix, H.W., Digmann, A., Jørgensen, P. & Møller Pedersen, K. 2012. *Hospitalsledelse: organisatorisk fænomen og faglig disciplin*, 2nd edition. Copenhagen: Gyldendal Business.

Bevan, G. & Hood, C. 2006. What's measured is what matters: Targets and gaming in the English public health care system, *Public Administration*, 84: 517-538.

Biørn, E., Hagen, T.P., Iversen, T. & Magnussen, J. 2003. The effect of activity-based financing on hospital efficiency: a panel data analysis of DEA efficiency scores 1992–2000. *Health Care Management Science*, *6*(4), 271-283.

Brunsson, N. 2002. *The organization of hypocrisy: Talk, decisions and actions in organizations*. 2nd edition. Oslo: Abstrakt.

Busse, R., Schreyögg, J. & Smith, P.C. 2006. Editorial: Hospital case payment systems in Europe, *Health Care Management Science,* 9, 211.

Christensen, J.G., Christiansen, P.M. & Ibsen, M. 2011. *Politik og forvaltning*. 3rd edition, Aarhus: Academica.

Christensen, J.G. & Pallesen, T. 2001. Institutions, distributional concerns, and public sector reform, *European Journal of Political Research,* 39(2), 179-202.

Christiansen, P. 1996. *Administrative funktioner på sygehuse*. Grenaa: Dansk Kommunalkursus.

Cots, F., Chiarello, P., Salvador, X., Castells, X. & Quentin, W. 2011. DRG-based hospital payment: Intended and unintended consequences. In Busse et al. (eds), *Diagnosis-related groups in Europe*: 75-92. UK: McGraw-Hill Education; Open University Press.

Damgaard, J.B. & Pallesen, T. 1998. Kommunalt selvstyre og organisering. In Blom-Hansen (ed.), *Offentligt og effektivt?: Institutionelle valg i den offentlige sektor*: 146-184. Kbh.: Gyldendal.

Davies, I.C. 1999. Evaluation and performance management in government, *Evaluation* 5 (2), 150-159.

De Araújo, J. & Branco, J.F.A. 2009. Implementing performance-based management in the traditional bureaucracy of Portugal', *Public Administration,* 87(3), 557-573.

DiMaggio, P. & Powell, W.W. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields, *American Sociological Review,* 48 (2), 147-160.

Dixit, A. 2002. Incentives and organizations in the public sector: An interpretative review, *The Journal of Human Resources,* 37(4), 696-727.

Donahue, A.K., Selden, S.C. & Ingraham, P.W. 2000. Measuring government management capacity: A comparative analysis of city human resources management systems, *Journal of Public Administration Research and Theory,* 10(2), 381-412.

Draborg, E., Kjær, T., Bech, M. & Mollerup, M. 2009*. Dokumentationsrapport – Spørgeskemaundersøgelse – PaRIS – Patientens Rejse i Sundhedssektoren.* Health Economics Papers, no. 5, University of Southern Denmark.

Dørken, R., Feilberg, S., Buse, T., Jakobsen, M., Pallesen, T., Hansen, X., Kristensen, S. & Bech, M. 2012. *Takster i faste rammer. En analyse af takst- og økonomistyringen på sygehusområdet*. Århus : KREVI.

Gilardi, F., Füglister, K. & Luyet, S. 2009. Learning from others, *Comparative Political Studies,* 42(4), 549-573.

Goldfinch, S. 2006. Rituals of reform, policy transfer, and the National University Corporation Reforms of Japan, *Governance,* 19(4), 585-604.

Hood, C. 2002. Control, bargains, and cheating: The politics of public-service reform, *Journal of Public Administration Research and Theory*, 12, 3, 309-332.

Hood, C. & Peters, G. 2004. The middle aging of new public management: Into the age of paradox?, *Journal of public administration research and theory*,*14*(3), 267-282.

H:S, Amtsrådsforeningen, Sundhedsstyrelsen, Finansministeriet & Indenrigs- og Sundhedsministeriet. 2005. *Evaluering af takststyring på sygehusområdet.* København: Indenrigs- og Sundhedsministeriet.

Indenrigs- og Sundhedsministeriet. 2005. *Produktivitet på sundhedsområdet: Åbenhed og fokus*. Velfærdstema. Indenrigs- og Sundhedsministeriet.

Jakobsen, M.L.F. 2009. Datarapport om undersøgelse af økonomistyringen på danske sygehuse og sygehusafdelinger i perioden 2005-2007. Aarhus: Institut for Statskundskab.

Jakobsen, M.L.F. 2010. The effects of new public management: Activity-based reimbursement and efficiency in the Scandinavian hospital sectors, *Scandinavian Political Studies,* 33(2), 113-134.

Jegers, M., Kesteloot, K., De Graeve, D. & Gilles, W. 2002. A typology for provider payment systems in health care, *Health Policy,* 60(3), 255-273.

Jespersen, P.K., Nielsen, L.-L.M. & Sognstrup, H. 2002. Professions, institutional dynamics, and new public management in the Danish hospital field, *International Journal of Public Administration,* 12, 1555-1574.

Lega, F. & Vendramini, E. 2008. Budgeting and performance management in the Italian National Health System (INHS) – Assessment and constructive criticism, *Journal of Health Organization and Management*, Vol. 22 No. 1, 2008, 11-22.

Meier, K.J. & O'Toole, L.J. 2009. The proverbs of new public management lessons from an evidence-based research agenda, *The American Review of Public Administration*, 39(1), 4-22.

Meyer, J.W. & Rowan, B. 1977. Institutionalized organizations: Formal structure as myth and ceremony, *The American Journal of Sociology,* 83(2), 340-363.

Miles, M.B. & Huberman, A.M. 1994. *Qualitative data analysis : An expanded sourcebook.* 2nd edition. (Thousand Oaks, CA: Sage.

Ministry of Finance. 2003. *Amternes økonomi 2004.* Copenhagen: Ministry of Finance.

Møller Pedersen, K., Bech, M. & Hansen, M.B. 2006. *Incitamentsstyring i sygehusvæsenet : Virkningen af øget takststyring og tilhørende incitamenter*. Odense: Syddansk Universitetsforlag.

Moynihan, D.P. 2005. Why and how do state governments adopt and implement »Managing for Results« reforms?, *Journal of Public Administration Research and Theory,* 15(2), 219-243.

Moynihan, D.P. 2008. *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

Nielsen, P.A. 2013. Performance management, managerial authority, and public service performance, *Journal of Public Administration Research and Theory*, 24(2), 431-458

Oliver, C. 1991. Strategic responses to institutional processes, *The Academy of Management Review,* 16(1), 145-179.

Or, Z. & Häkkinen, U. 2011. DRGs and quality: For better or worse. In Busse et al. (eds), *Diagnosis-related groups in Europe*: 115-130. UK: McGraw-Hill Education, Open University Press.

O'Reilly, J., Busse, R., Häkkinen, U., Or, Z., Street, A. & Wiley, M. 2012. Paying for hospital care: The experience with implementing activity-based funding in five European countries, *Health economics, policy, and law*. 7:73.

Ouchi, W.G. 1980. Markets, bureaucracies, and clans, *Administrative Science Quarterly,* 25(1), 129-141.

Pettersen, I.J. & Nyland, K. 2006. Management and control of public hospitals: The use of performance measures in Norwegian hospitals. A case-study, *The International Journal of Health Planning and Management*, 21(2), 133-149.

Pettersen, I.J. 1999. Accountable management reforms: Why the Norwegian hospital reform experiment got lost in implementation, *Financial Accountability & Management,* 15(3-4), 377-396.

Pollitt, C. & Bouckaert, G. 2011. *Public management reform: A comparative analysis – New public management, governance, and the neo-weberian state*. Oxford University Press.

Propper, C. & Wilson, D. 2003. The use and usefulness of performance measures in the public sector, *Oxford review of economic policy* 19 (2), 250-267.

Robinson, M. 2002. *Output purchase funding and budgeting systems in the public sector*. Vol. 22. Blackwell Publishers Inc.

Robinson, M. 2007a. Performance budgeting models and mechanisms. In Robinson (ed.), *Public sector performance budgeting – Linking funding and results*: 1-20. Washington: International Monetary Foundation.

Robinson, M. 2007b. Making performance budgeting work. In Robinson (ed.), *Public sector performance budgeting – Linking funding and results*: 113-127. Washington: International Monetary Foundation.

Robinson, M. 2007c. Purchaser-provider system. In Robinson (ed.), *Public sector performance budgeting – Linking funding and results*: 296-311. Washington: International Monetary Foundation.

Scott, W.R. & Davis, G.F. 2007. *Organizations and organizing: Rational, natural and open systems perspectives*. International edition. Harlow: Pearson / Prentice Hall.

Serritzlew, S. 2006. Linking budgets to activity: A test of the effect of output-purchase budgeting, *Public Budgeting & Finance*, Summer, 101-120.

Smith, P.C. 2007. Formula funding and performance budgeting. In Robinson (ed.), *Public sector performance budgeting – Linking funding and results*: 211-233. Washington: International Monetary Foundation.

Strandberg-Larsen, M., Nielsen, M., Valgårda, S., Krasnik, A. & Vrangbæk, K. 2007. Denmark – Health system review, *Health Systems in Transition*, 9(6).

Street, A., O'Reilly, J., Ward, P. & Mason, A. 2011. DRG-based hospital payment and efficiency: Theory, evidence, and challenges. In Busse et al. (eds), *Diagnosis-related groups in Europe*: 93-114. UK: McGraw-Hill Education; Open University Press.

Utz, H. 2010. Better performance with performance budgeting? Analyzing cases of success and failure in public administrations, *International Public Management Journal,* 13(4), 408-428.

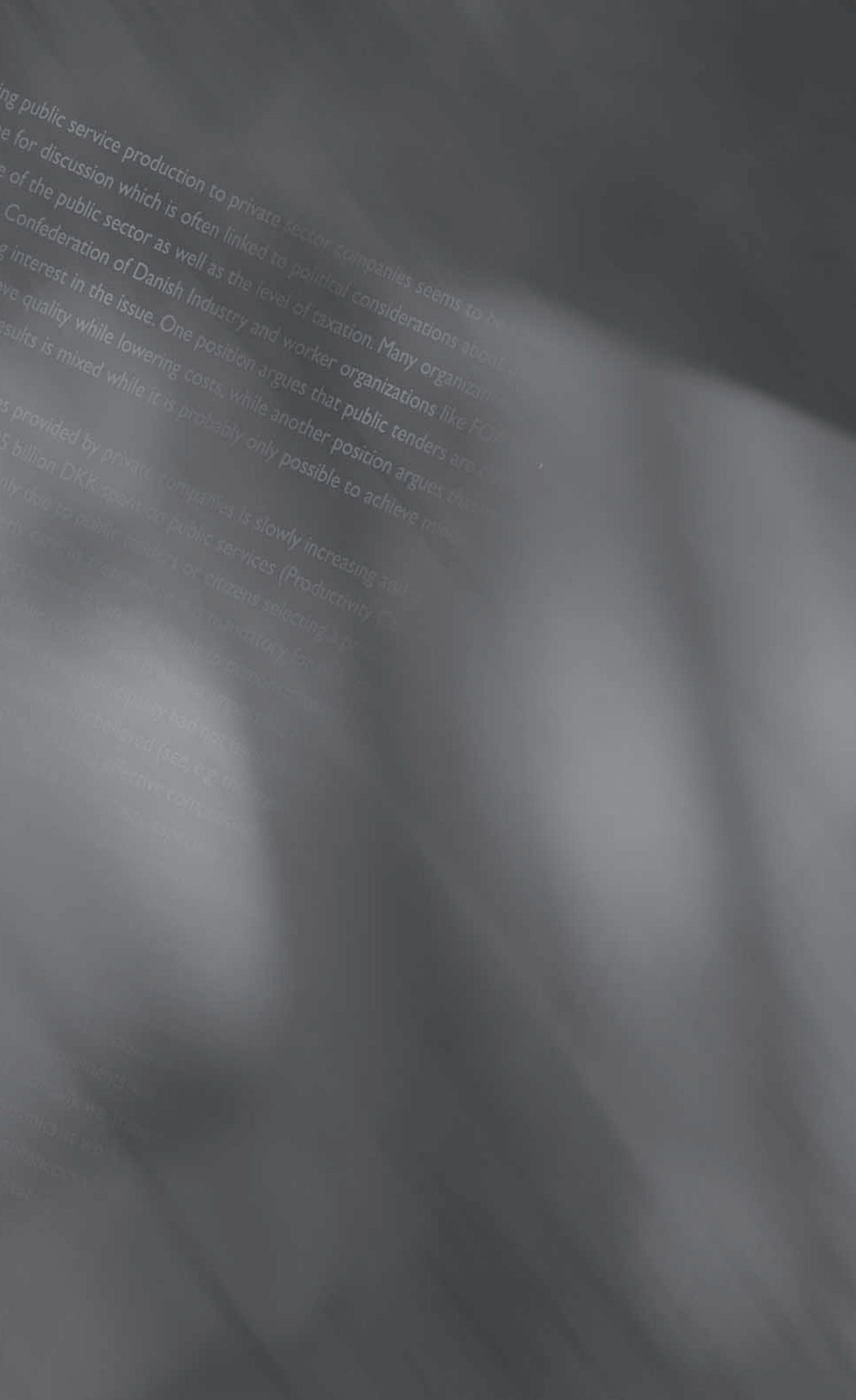
Wiley, M. 2011. From the origins of DRGs to their implementation in Europe. In Busse et al. (eds), *Diagnosis-related groups in Europe*: 3-8. UK: McGraw-Hill Education; Open University Press.

Vrangbæk, K. 2012. The Danish health care system, 2011. In Thomson et al., *International profiles of health care systems, 2011*: 31-37. The Commonwealth Fund.

Zucker, L.G. 1987. Institutional theories of organization, *Annual Review of Sociology,* 13, 443-464.

Notes

1. This includes two hospitals from each region except Region Zealand, where only one hospital wanted to participate.
2. The subsequent results also hold if we only look at the rate of reimbursement before and after the baseline, respectively.
3. Test of differences in means between two groups.
4. Test of difference in proportions between two groups.
5. Test of difference in proportions between two groups.
6. Test of difference in proportions between two groups; it is significant, however, at an 80 percent confidence level.
7. Based on a two-sample Wilcoxon rank-sum (Mann-Whitney) test that fits the ordinal nature of the data. Performing a standard test of differences in group means also shows a significant difference; $p = 0.0053$.



Editors' corner

Outsourcing service production in the public sector: Are we addressing the right question?

By Per Nikolaj Bukh, Kurt Klaudi Klausen, Dana Minbaeva, Niels Peter Mols and Flemming Poulfelt

Outsourcing public service production to private sector companies seems to be a never-ending theme for discussion which is often linked to political considerations about the appropriate size of the public sector as well as the level of taxation. Many organizations including, e.g. the Confederation of Danish Industry and worker organizations like FOA have, of course, a strong interest in the issue. One position argues that public tenders are clearly beneficial and improve quality while lowering costs, while another position argues that evidence regarding the results is mixed while it is probably only possible to achieve minor cost reductions.

The share of public services provided by private companies is slowly increasing and is today around 25 % of the 385 billion DKK spent on public services (Productivity Commission (2014b, p. 148). This is mainly due to public tenders or citizens selecting a private provider for a public service, e.g. elderly care, in areas where it is mandatory for municipalities to offer citizens a choice of service providers. It can be difficult to demonstrate how price and quality are actually influenced by public tenders as it is almost impossible to determine what the price and quality would have been, if the municipality had not issued a call for tenders (c.f. Kristensen 2014). However, it is generally believed (see, e.g. the argument put forward by the Productivity Commission, 2014a) that: (a) effective competition in the area of public services increases productivity as it leads to lower costs, especially when economies-of-scale are prevalent; (b) improved cooperation between the private and public sectors increases innovation and; (c) increased use of specialized firms improves managerial focus, thus leading to more effective service provision. Our aim with this editorial is *not* to argue for more or less outsourcing of public services. Rather we suggest that the decision to engage private companies or not to deliver services should be based on a concrete analysis of the objectives of the outsourcing decision (or more generally the make-or-buy decision in economic theory), a better understanding of how the improvements will be realized and what is expected from the private-public partnership. Specifically, we argue that

more attention should be directed towards understanding the dynamics at work when private companies and the public sector deliver the same services simultaneously. This phenomenon, known as *concurrent sourcing,* is widespread in both the private and public sector, but seems to have been overlooked in discussions about outsourcing service production in the public sector.

From competition to productivity

The benefits resulting from improved competition are frequently mentioned as the main argument for outsourcing public services: The more competitors and the fewer barriers to competition, the higher productivity, the better market forces function and the lower the prices. This basic message of economics 101 has been well supported empirically and, perhaps not surprisingly, also permeates the recommendation of the Productivity Commission (2014b). If market forces are to function, we of course need to have a market; and no market exists without potential for exchange. Thus, it follows that we need private companies to deliver public services in order to reap the benefits of competition.

However, the theory does not tell us exactly how many suppliers are needed, or what share of a specific service should be outsourced to obtain most of the benefits from increased competition. Further, what does outsourcing a certain percentage of services to the private sector imply for the functioning of the market forces? If we assume that no specific share of outsourcing is optimal, it follows that politicians should avoid setting targets for outsourcing. It may be that the "optimal" level of outsourcing depends on how the service delivery system is structured in, e.g. a specific municipality, while factors inherent in the types of service may be important.

The make or buy decision – is it an either-or question?

The question implied by the discussion above is whether an organization should buy specific components or services or whether it should produce them internally? This is a question all public sector organizations have to answer. Private firms face a similar problem. However, the question whether to make *or* buy may be the wrong question.

With the term *concurrent sourcing* researchers have recently focused on an alternative solution. Concurrent sourcing refers to "backward, partial vertical integration of a homogeneous good (or service) by a single firm" (Parmigiani, 2007, p. 285). So concurrent sourcing occurs when a firm or public organization buys a good or service from an external supplier while simultaneously producing the same good or service internally. For example, a Danish municipality engages in concurrent sourcing when it has an internal unit that maintains its roads, while simultaneously contracting a private firm to also conduct road maintenance. Similarly, concurrent sourcing occurs when one or more nursing homes in a municipality are operated by private companies while the remainder of the nursing homes is run by the municipality.

When we focus on the provision of basic welfare services such as care for the elderly, primary schools and various social services as well as hospital services provided at the regional level, it can generally be observed that responsibility for service provision is almost never outsourced to private sector companies. When private hospitals deliver services in specific areas, we almost always see that similar services are being delivered by public hospitals simultaneously. Also, with respect to primary schools, to our knowledge, no municipality has considered relinquishing responsibility for the public school system. Further, most municipalities prefer to have a mixture of social services, some of which are bought from private companies, some from other municipalities while others are produced internally. The question is whether this is a specific public sector phenomenon and how can it be explained?

Research on concurrent sourcing

A brief review of the literature on concurrent sourcing allows us to identify some important conclusions. *First*, concurrent sourcing is a widespread phenomenon. For example, Parmigiani (2007) focuses on the sourcing decisions of metal stamping and powder metal firms for production tooling and services, and she finds that 28% of the outsourcing decisions in her sample resulted in concurrent sourcing. Heide *et al.* (2014) find similar results among clothing manufacturers where 30 % of the firms used concurrent sourcing. In Scandinavia, case studies from the wood product manufacturing industry confirm that concurrent sourcing is widely used (Nordigården *et al.,* 2014). A few studies also show the significant use of concurrent sourcing in the public sector (e.g. Hefetz *et al.,* 2014).

Second, empirical results suggest that concurrent sourcing is a stable, equilibrium sourcing mode (e.g. Parmigiani, 2007; Heide *et al.,* 2014). Thus, concurrent sourcing is not merely a temporary phenomenon that can be observed when firms are in the process of outsourcing or insourcing production. Furthermore, concurrent sourcing is a unique sourcing mode that is empirically different from other sourcing arrangements such as long-term contracts between buyers and suppliers, and it also has attributes which are different from the average attributes of internal production and external suppliers (Parmigiani, 2007). In other words, there may be important synergies between internal production and the use of external suppliers.

Third, economic models suggest that the total cost of sourcing may be lower when internal production is combined with external suppliers, and that the cost depends on the relative quantities produced internally and sourced from external suppliers (e.g. Puranam *et al.,* 2013). Therefore, firms and public organizations should ask the question: how much of a certain good or service should we produce ourselves and how much should we purchase?

Can theory explain practice?

The decision to produce internally or to purchase from external suppliers depends on the relative transaction and production costs of the two solutions and there are several theoretical perspectives that attempt to explain this choice. Three of the most prominent theoretical perspectives are transaction cost theory, resource-based theory, and neoclassical economics.

Neoclassical economics explains the choice between to produce or to purchase by the ability of a firm to operate internally at efficient production scales. If a firm is able to produce efficiently in-house, then the solution is to produce in-house. However, if small scale production makes in-house production inefficient, then the firm should use external suppliers. External suppliers can supply more customers thereby exploiting economies of scale in the form of lower production costs.

Transaction cost theory assumes that production costs are lower when a firm uses external suppliers. However, the use of external suppliers may result in lower product and service quality, while they may refuse to adapt, or may raise prices or in other ways act opportunistically when they have the power to do so. This leads to higher transaction costs. In order to avoid opportunistic suppliers, firms choose long-term contracts with external suppliers or they internalize production.

Resource-based theory does not assume that production costs are lower when firms use external suppliers. Instead, the theory suggests that the choice between internal production and external suppliers depends on who possesses superior production capabilities and resources. If the external supplier has superior production capabilities and hence low production costs, then the firm in question should source from the external supplier. On the other hand, superior internal resources and capabilities make in-house production efficient.

However, these three perspectives do not immediately explain concurrent sourcing. They answer the produce-or-purchase question, but do not consider concurrent sourcing and the decision regarding how much to purchase and how much to produce. Therefore, the perspectives need to be supplemented by economic explanations for concurrent sourcing.

Towards new theories

There may be several reasons why concurrent sourcing has lower total costs than in-house production or external suppliers (cf. Nordigården et al., 2014). With concurrent sourcing it is possible to use information from in-house production to evaluate and control external suppliers and vice versa. A small in-house production also demonstrates the ability to further integrate and hence replace external suppliers with in-house production. This positive effect of concurrent sourcing has been empirically supported by Heide *et al.* (2014) who find that, in a concurrent sourcing context, supplier opportunism towards purchasers is reduced when purchasers monitor suppliers, whereas monitoring in a singular sourcing context is less effective.

In-house production also increases a purchaser's ability to transfer knowledge to and receive knowledge from the external suppliers. Therefore, concurrent sourcing makes it easier for the buyer and supplier to learn from each other and thereby take advantage of both internal and external resources and capabilities. Since a firm facing technological uncertainty often needs a broad technological knowledge base, it is more likely to use both internal and external sources to maintain and develop this knowledge base. Therefore, technological uncertainty increases the likelihood of concurrent sourcing (Parmigiani, 2007).

Finally, with unpredictable demand, firms sometimes face demand that exceeds their internal production capacity. In such cases, they may choose to expand capacity or use existing capacity more intensively. However, both solutions raise production costs and may not be technologically feasible in the short term. With external suppliers willing to satisfy varying demand, concurrent sourcing can be used if a firm has too little internal capacity, while internal excess capacity may be avoided.

Is concurrent sourcing the solution?

Concurrent sourcing has many advantages, but these advantages do not come without cost. For example, it can be costly to set up and manage both in-house production and external suppliers. The division of production among internal and external production facilities may make it impossible to operate at an efficient scale, while the addition of in-house production may damage relations with an external supplier.

It is difficult to determine exactly when public organizations and firms should use concurrent sourcing. The literature has identified a number of advantages and disadvantages of concurrent sourcing compared with alternative sourcing modes. Some of the key benefits of concurrent sourcing include an improved ability to monitor suppliers due to reduced information asymmetries, increased learning due to the combination of knowledge gained from in-house production and more diverse knowledge from external sources as well as protection against supplier opportunism (cf. Parmigiani 2007) – especially in the presence of performance uncertainty (Dutta *et al.,* 1995), and information asymmetries (Heide, 2003).

Such insights give us an idea as to when the synergies and other advantages connected with concurrent sourcing make it more efficient than alternative sourcing modes. If we, for instance, compare home care according to the Act on Social Services (Serviceloven, §83) with more complex rehabilitation services, both performance uncertainty and informational asymmetries are larger in the latter case as is the learning potential from combining general knowledge with deep tacit domain-specific knowledge, thereby increasing the likelihood that the benefits of concurrent sourcing outweigh, e.g. economics of scale and a higher degree of outsourcing.

However, this does not provide managers with concrete guidance about when to choose concurrent sourcing, and the choice is further complicated by the many possible types of concurrent sourcing that exist. Besides deciding how much to buy and how much to make, firms also have to decide what kind of contract and relationship to develop with the external supplier. Without understanding the dynamics of and synergies between in-house production and external suppliers in a specific service area, it may be impossible to answer whether public or private firms should produce public services. Consequently, we propose that more effort should be put in understanding the many alternative sourcing modes theoretically as well as practically. Additional reports which survey the use or otherwise of private suppliers, or which calculate potential based on more or less unrealistic assumptions will probably not be able to answer how the production of public services is best organized.

In light of the fact that concurrent sourcing is widely adopted, it is surprising that we seem to have little knowledge about the advantages of different types of concurrent sourcing, and thus why concurrent sourcing may be more efficient than alternative modes of sourcing. Hopefully, greater insight into this phenomenon will be obtained as more studies focus on explaining the advantages and disadvantages of alternative sourcing arrangements in the public sector.

References

Dutta S, Bergen M, Heide JB, John G. 1995. Under-standing dual distribution: the case of reps and house accounts, *Journal of Law, Economics, and Organization* 11:189–204.

Hefetz, A., Warner M. & Vigoda-Gadot, E. 2014. Concurrent sourcing in the public sector: A strategy to manage contracting risk, *International Public Management Journal* 17(3): 365-386.

Heide J.B. 2003. Plural governance in industrial purchasing. *Journal of Marketing* 67:18–29.

Heide, J.B., Kumar, A. & Wathne K.H. 2014. Concurrent sourcing, governance mechanisms, and performance outcomes in industrial value chains, *Strategic Management Journal* 35(8):1164-1185.

Kristensen, N. 2014. Økonomisk effektivitet blandt private og offentlige udbydere af ældrepleje: Metodiske overvejelser og empiriske resultater. KORA, The National Institute of Local and Regional Government Analysis and Research, March 2014.

Nordigården, D., Rehme, J., Brege, S., Chicksand, D. & Walker H. 2014. Outsourcing decisions – the case of parallel production, *International Journal of Operations & Production Management* 34(8):974-1002.

Parmigiani, A. 2007. Why do firms both make and buy? An investigation of concurrent sourcing, *Strategic Management Journal* 28(3):285-311.

Productivity Commission. 2014a. Offentlig-Privat Samarbejde, Produktivitetskommissionen, Analyserapport 6, februar 2014.

Productivity Commission. 2014b. Det handler om velstand og velfærd, Produktivitetskommissionen, Slutrapport, marts 2014.

Puranam, P., Gulati, R. & Bhattacharya, S. 2013. How much to make and how much to buy? An analysis of plural sourcing strategies?, *Strategic Management Journal* 34(10):1145-1161.

PUBLISHER
DJØF Publishing
137 Gothersgade
DK-1123 Copenhagen
Tel: +45 39135500
Email: forlag@djoef.dk
www.djoef-forlag.dk

SUBSCRIPTION
Printed version
Normal subscription:
Dkk. 718.75 pr. year (VAT incl.)
Members of DJØF:
Dkk. 328.75 pr. year (VAT incl.)
Students:
Dkk. 178.75 pr. year (VAT incl.)
Postage not included in subscription prices
Digital version
Dkk. 950.00 pr. year (VAT incl.) with free access to the archive of articles.

Each volume has four numbers pr. year. To be published in February, June, September and December

Printing: Ecograf, Denmark

ISSN 0902-3704

Aim and scope of the Journal
The *Danish Journal of Management & Business* is a generalist, academic journal covering all management fields. The *DJMB* aims to present the latest thinking and research on major management topics with relevance for a Danish readership. While meeting high academic quality standards, articles are interesting and accessible to non-specialists.

Articles should be an accessible source of useful information and ideas for practitioners and students. The Journal takes a broad view of business and management and encourages submissions from all disciplines if they contribute significantly to problems considered by Danish managers and specialists. The *DJMB* is a peer-reviewed journal, with a 'double-blind' procedure involving at least two reviewers.

The key focus of DJMB will be on conceptual innovation and practical application to enhance the field of management thinking, practice, and research. The main objective of the journal is to provide a communication medium for all those working in the management field whether from industry, consulting or academic institutions. Contributions will therefore be welcomed from practitioners, academics and professionals working in private, public and non-profit organizations.

The DJM is distributed free of charge to about 12.000 members of DJØF, including student members, and to individual and institutional subscribers. A recent analysis of the readership indicate that articles are read by up to 4.000 individuals.

Editorial criteria
Articles should take a conceptual or theoretical point of departure and include practical application. Papers will typically be based on research, report a case study situation, report results from survey, or any combination thereof. The prime requirement for acceptance of an article will be that it

1. Provide new information relevant to the readership
2. Makes a significant original contribution to the field
3. Is directly relevant to for practice or society
4. Bridge the gap between theory and practice
5. Report results or theory relevant for Danish firms
6. Contains elements which have general application
7. Is within the scope of the journal coverage.

Every year a best-paper prize of DKK 40.000 is awarded to celebrate and applaud a well-written paper that combines academic rigor and validity with practical relevance.

The editorial process
Any article submitted to DJMB should be prepared according to the style guide for authors, but should omit any author-identifying references from the title page. Please submit the title page with authors name, title and affiliation as a separate file. Manuscripts should be submitted by email to one of the Editors.

Manuscripts submitted will be judged primarily on their substantive content, but writing style, structure, and length are also considered.

Publication decisions by the Editors are facilitated by reviews by qualified reviewers using a 'double-blind' review process. Reviewers' comments are made available to authors. Authors are strongly encouraged to seek peer review prior to submission. Also potential contributors are encouraged to contact the Editors in advance to seek advice if they are in doubt about the appropriateness of the manuscript for DJMB. Manuscripts that are inappropriate or insufficiently developed will be returned to the authors without formal review.

The review process follows international standards and the outcome can be either that the manuscript is accepted for publication in its present form, that revisions are suggested or that the manuscript is rejected. The editors will advise the author(s) how to revise the manuscript in a way that makes the paper appropriate for the readerships of the journal.

Style guide
For further information on guidelines for authors and the editorial process visit: http://www.djoef-forlag.dk/da/tidsskrifter/djm

Danish Journal of Management & Business
DJØF Publishing
137 Gothersgade
DK-1123 Copenhagen

Editorial board

Authors

Dag Øivind Madsen
Assistant professor, ph. d.
Buskerud and Vestfold University College
Bredalsveien 14
Hønefoss
NO-3511 Buskerud
Norway
e-mail: Dag.Oivind.Madsen@hbv.no

Steen Nielsen
Associate Professor, ph.d
School of Business and Social Sciences
Department of Economics and Business
University of Aarhus
Fuglesangs Allé 4
DK-8210 Aarhus V
e-mail: sni@asb.dk

Erland H. Nielsen
Associate Professor
School of Business and Social Sciences
Department of Economics and Business
University of Aarhus
Fuglesangs Allé 4
DK-8210 Aarhus V
e-mail: ehn@asb.dk

Anders Jacobsen
Commercial Analyst
Project Trition
Siemens Wind Power A/S
Borupvej 16
DK-7330 Brande
e-mail: anders.jacobsen@siemens.com

Lars Bjørn Pedersen
Controller
Pindstrup Mosebrug A/S
Fabriksvej 2, Pindstrup
DK-8550 Ryomgaard
e-mail: larsbjornpedersen@gmail.com

Peter Dahler-Larsen
Professor, ph.d.
Department of Political Science
University of Copenhagen
Øster Farimagsgade 5
Postbox 2099
DK-1014 Copenhagen K
e-mail: pdl@ifs.ku.dk

Signe Pihl-Thingvad
Associate professor
Department of Political Science and Public Management
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
e-mail: ssp@sam.sdu.dk

Bente Bjørnholt
Senior Researcher
The Danish National Centre for Social Research
Købmagergade 22
DK-8200 Aarhus N
e-mail: BeBj@kora.dk

Andrej Christian Lindholst
Assistant professor
Department of Political Science
Aalborg University
Fibigerstræde 1
DK-9220 Aalborg Øst
e-mail: acl@dps.aau.dk

Jeppe Agger Nielsen
Assistant professor
Department of Political Science
Aalborg University
Fibigerstræde 1
DK-9220 Aalborg Øst
e-mail: agger@dps.aau.dk

Mads Leth Felsager Jakobsen
Associate professor, ph.d.
Department of Political Science and Government
University of Aarhus
Bartholins Allé 7
DK-8000 Aarhus C
e-mail: mads@ps.au.dk